

CLINICAL RESEARCH CENTER  
Karolinska Institutet, Stockholm, Sweden

Structural and genetic  
characterization of  
lipopolysaccharides extracted from  
disease causing non-typeable  
*Haemophilus influenzae* strains

Mikael Engskog



**Karolinska  
Institutet**

Stockholm 2010

All previously published papers were reproduced with permission from the publisher.

Published by Karolinska Institutet. Printed by Larserics Digital Print AB

© Mikael Engskog, 2010  
ISBN 978-91-7409-731-3

*To my beautiful son and lovely wife, stars of my life*



## ABSTRACT

---

This thesis deals with the structure and genetic blueprint of lipopolysaccharides (LPS) expressed by the Gram-negative bacterium *Haemophilus influenzae* (*H. influenzae*).

*H. influenzae* is an opportunistic pathogen that regularly colonizes the upper respiratory tract and exists in encapsulated (typeable) or nonencapsulated (non-typeable) forms (NTHi). Prior research has indicated that the surface expressed lipopolysaccharides (LPS) is a major virulence factor of *H. influenzae*. Pathogenic behavior can for example result in respiratory tract infections, otitis media (OM) or invasive disease such as meningitis.

The thesis contains detailed studies of the phase-variable glycosyltransferase *lex2*. This transferase was shown to act either as a glucosyltransferase or galactosyltransferase depending on one single key amino acid. This was established by elucidation of LPS expressed by genetically defined *lex2* mutant strains or transformant strains in which the transferase activity was removed or altered.

Moreover, two structural characterizations of non-typeable clinical isolates, strains 2019 and R2866 are investigated. NTHi R2866 is an atypical non-typeable strain as it was isolated from a child with meningitis. Our data indicate that R2866 produces an extremely heterogeneous population of glycoforms with expression of L-glycero-D-manno-heptose (L,D-Hep) in its outer core. This residue was evidenced to carry a phosphocholine (PCho) residue in O-7 position, a substitution which is novel. Moreover, this study includes data obtained for genetically defined mutant strains R2866*lpsA* and R2866*losB2* as these demonstrated detailed LPS structures not seen in wild-type.

The characterization of LPS expressed by 2019 wild-type strain and mutant strains 2019*lex2*, 2019*lpt3* and 2019*pgmB* add to the previously published structure of 2019. In 1992 it was established that lactose is linked to the proximal heptose (Hep I) of the conserved triheptosyl inner-core moiety, no other structures were reported. We show that the middle heptose (Hep II) can express a  $\beta$ -D-Galp-(1 $\rightarrow$ 4)- $\beta$ -D-Glcp-(1 $\rightarrow$ 4)- $\alpha$ -D-Glcp-(1 $\rightarrow$ 3) epitope. Interestingly, the *lex2* mutant was indicated to be substituted at O-2 at Hep III by  $\beta$ -D-Glcp which, in turn, can be further extended. Such elongations have never been reported for NTHi 2019 before. NTHi 2019 belongs to a small subset of strains that express one additional phosphoethanol amine (PEtn) residue in its outer core. Here, we establish that PEtn substitutes O-3 of the distal heptose (Hep III). This PEtn substituent was absent in the *lpt3* mutant indicating Lpt3 to be the transferase required to add PEtn to Hep III.

## LIST OF PUBLICATIONS

---

This thesis is based on the following papers, which in the text will be referred to by their roman numerals (I-IV).

- I. **Engskog MKR**, Yildirim HH, Li J, Richards JC, Deadman M, Hood DW and Schweda EKH.  
A dual role for the *lex2* locus: identification of galactosyltransferase activity in non-typeable *Haemophilus influenzae* strains 1124 and 2019.  
*Carbohydrate Research*. 344(5): 632-641, (2009)
- II. Deadman ME, Hermant P, **Engskog MKR**, Makepeace K, Moxon ER, Schweda EKH and Hood DW.  
*Lex2B*, a phase-variable glycosyltransferase, adds either a glucose or a galactose to *Haemophilus influenzae* lipopolysaccharide.  
*Infection and Immunity*. 77(6): 2376-2387, (2009)
- III. **Engskog MKR**, Deadman ME, Hood DW and Schweda EKH.  
Detailed structural features of lipopolysaccharide glycoforms in *NTHi* strain 2019  
*Manuscript*
- IV. **Engskog MKR**, Deadman ME, Li J, Hood DW and Schweda EKH.  
Structural analysis of the lipopolysaccharide from invasive non-typeable *Haemophilus influenzae* strain R2866  
*Manuscript*

## LIST OF ABBREVIATIONS

---

1D	One-dimensional
2D	Two-dimensional
Ac	Acetate
CE	Capillary electrophoresis
CID	Collision induced dissociation
COSY	Correlation spectroscopy
CRP	C-reactive protein
DMSO	Dimethyl sulfoxide
ESI	Electrospray ionization
Gal	Galactose
GalNAc	2-acetamido-2-deoxygalactose
GC	Gas chromatography
Glc	Glucose
GlcNAc	2-acetamido-2-deoxyglucose
Gly	Glycine
GPC	Gel-permeation chromatography
Hep	Heptose
D,D-Hep	<i>D-glycero-D-manno</i> -heptose
L,D-Hep	<i>L-glycero-D-manno</i> -heptose
Hex	Hexose
HexNAc	<i>N</i> -acetylhexoseamine
Hib	<i>Haemophilus influenzae</i> type b
HMQC	Heteronuclear multiple quantum coherence
HPAEC	High performance anion exchange chromatography
HPLC	High performance liquid chromatography
Kdo	3-deoxy- <i>D-manno</i> -oct-2-ulsonic acid
AnKdo-ol	Reduced anhydro Kdo
lipid A-OH	<i>O</i> -deacylated lipid A

## LIST OF ABBREVIATIONS

---

LPS	Lipopolysaccharide
LPS-OH	<i>O</i> -deacylated LPS
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MS <sup>n</sup>	Multiple step tandem mass spectrometry
NAD	Nicotinamide adenine dinucleotide
Neu5Ac	<i>N</i> -acetylneuraminic acid
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser enhancement
NOESY	NOE spectroscopy
NTHi	Non-typeable <i>Haemophilus influenzae</i>
OS	Oligosaccharide
OM	Otitis media
PAD	Pulsed amperometric detection
PAF	Platelet activating factor
<i>P</i>	Phosphate
<i>P</i> Cho	Phosphocholine
PCP	Phenol:chloroform:light petroleum
PCR	Polymerase chain reaction
<i>PE</i> tn	Phosphoethanolamine
PMAA	Partially methylated alditol acetate
<i>PP</i> Et <sub>n</sub>	Pyrosphosphoethanolamine
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SIM	Selective ion monitoring
TIC	Total ion chromatogram
TFA	Trifluoroacetic acid
TOCSY	Total correlation spectroscopy



## TABLE OF CONTENTS

---

<b>INTRODUCTION</b>	1
<i>Haemophilus influenzae</i>	2
The fundamentals of <i>H. influenzae</i> LPS structure	4
Lipid A, structure and assembly	5
Transport of LPS to the outer membrane	7
The variable outer core	8
Host mimicry and phase variation	11
Sialic acid and implications for virulence	12
Non-carbohydrate residues	13
Colonization and infection	15
<b>MATERIALS AND METHODS</b>	17
<b>Bacterial cultivation and construction of mutant/transformant strains</b>	18
<b>Extraction of LPS</b>	18
<b>Preparation of material</b>	19
Hydrazinolysis of LPS	19
Hydrolysis of LPS and preparation of lipid A	19
Dephosphorylation and deacylation of OS	20
Compositional absolute configuration analysis	20
Methylation and permethylation analysis	22
<b>Mass spectrometry</b>	22
GC-MS	22
ESI-MS	23
HPLC-ESI-MS <sup>n</sup> and CE-ESI-MS/MS	24
<b>NMR spectroscopy</b>	25
1D <sup>1</sup> H NMR	26
2D homonuclear <sup>1</sup> H- <sup>1</sup> H correlation spectroscopy	27
2D heteronuclear <sup>1</sup> H- <sup>13</sup> C and <sup>1</sup> H- <sup>31</sup> P correlation spectroscopy	29

## TABLE OF CONTENTS

---

<b>HPAEC-PAD and GPC</b>	29
<b>AIM OF THE THESIS</b>	30
<b>RESULTS</b>	
<b>PAPER I</b>	31
<b>PAPER II</b>	33
<b>PAPER III</b>	35
<b>PAPER IV</b>	38
<b>CONCLUSSIONS AND FUTURE PERSPECTIVES</b>	40
<b>POPULÄRVETENSKAPLIG SAMMANFATTNING</b>	43
<b>ACKNOWLEDGEMENTS</b>	45
<b>REFERENCES</b>	47

## INTRODUCTION

---

### INTRODUCTION

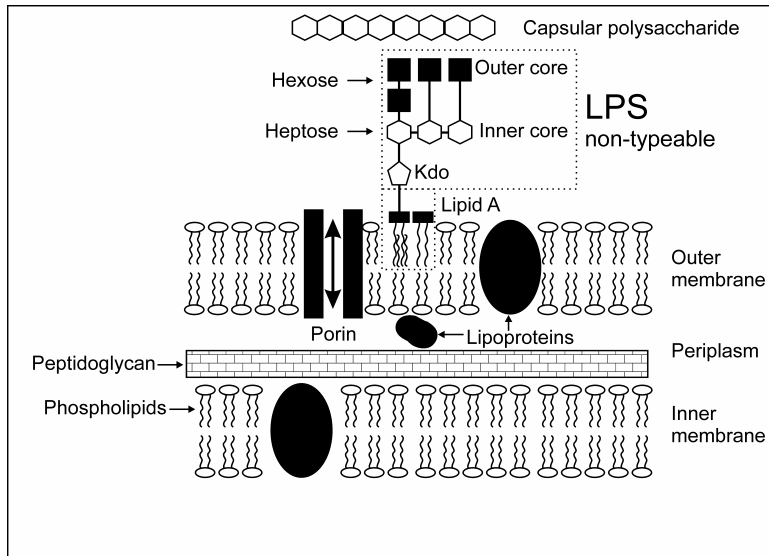
One of the most common reasons for children to meet their pediatrician during their first eventful year of life is middle ear infection, otitis media (OM) (Murphy TF, 2009). Such infections can be troublesome as they can reoccur frequently during early age. Having multiple middle ear infections as a child can in the long run affect the child's hearing capacity and ultimately result in language difficulties. Early infection typically involves the child obtaining a cold which later also affects the middle ear. Inflammation in the middle ear leads to pain which normally settles in a day or two, in rare cases it can last for a week. In more serious cases, the ear drum ruptures causing discharge of pus from the ear. The ruptured drum will usually heal quite fast, but reoccurring infections may lead to scarring of the drum that causes hearing disabilities. OM is caused primarily by three bacteria, *Streptococcus pneumoniae*, *Moraxella catarrhalis* and non-typeable *Haemophilus influenzae* (NTHi) of which the later is the focus of this thesis.

This thesis deals with the concept of lipopolysaccharides (LPS) that are found on the surface of *H. influenzae*. These carbohydrate-containing glycolipids play vital roles in both manifestation and pathogenic behavior of the bacterium. The work presented here is the structural elucidation of LPS and characterization of genes involved in the biosynthesis of LPS. It is believed that the research will lead to both better diagnostics and future treatments for patients infected by *H. influenzae*.

Bacteria are commonly divided into two groups based on the experiments performed by the Danish physician Hans Christian Gram during the later 1900<sup>th</sup> century. Gram's ambition was to distinguish between difficult-to-see bacteria that caused pneumonia and mammalian cell nuclei in infected tissue. To his disappointment, he noted that certain bacteria did retain the developed stain while others did not (Lim D, 1998). It was later established, much due to Gram's work, that two types of bacteria exist. Bacteria could be classified as being either Gram-positive (those that retain the stain) or Gram-negative (those that do not retain the stain, but could be stained with the counter stain safranin). The reason for the different abilities to retain the respective dyes lie in the composition of the cell wall.

Gram-positive bacteria have a thick cell wall consisting of a cytoplasmic membrane protected by a peptidoglycan-containing layer. The cell wall of Gram-negative bacteria is more defined and consists of a thin sheet of peptidoglycan structure positioned between two membranes called the outer and inner membrane, this area is referred to as the periplasm. The outer membrane consists of a phospholipid bilayer interspersed with proteins and lipoproteins. The main interest for this thesis is the LPS that decorates the surface of the outer membrane indicated in the schematic shown in **Figure 1**.

## INTRODUCTION



**Figure 1.** Schematic picture of the Gram-negative cell wall of *H. influenzae*.

### *Haemophilus influenzae*

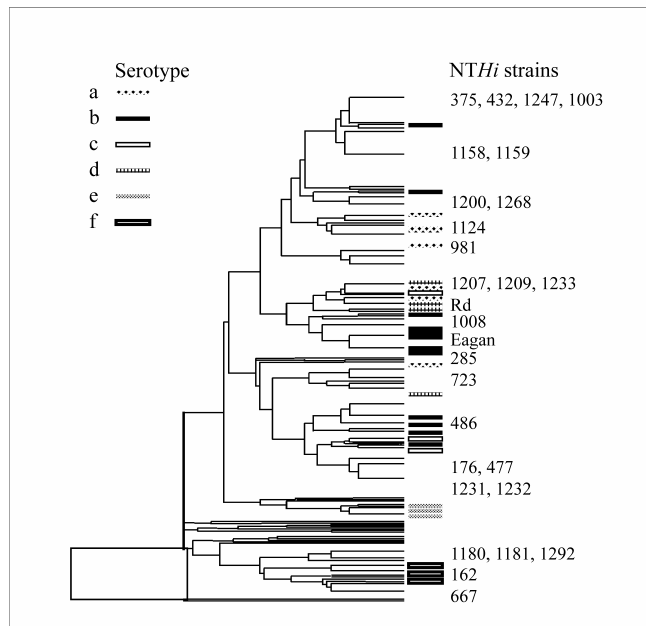
The generic name *Haemophilus* refers to the bacterium's need for blood in order to survive, as haemophilus literally means "blood-loving". Blood contains two growth factors, haemin and nicotinamide adenine dinucleotide (NAD) which both are essential for the bacterium to thrive within the host. The bacterium was first described by Dr. Robert Pfeiffer in 1892 during a raging epidemic influenza period. As he believed the newly discovered bacterium was linked to the onset of influenza, he falsely gave it the name *H. influenzae*.

*H. influenzae* is a small, fastidious coccobacillus that frequently colonizes the human respiratory tract acting either as commensal or pathogen (Erwin AL, 2007, Foxwell AR, 1998, Gilsdorf JR, 1998, Moxon ER, 1990a, Murphy TF, 2003). Pathogenic behavior can result in acute or chronic OM and/or lower respiratory tract infections (Campagnari AA, 1987, St Geme JW, 3rd, 2000). In more severe cases, patients with typeable *H. influenzae* can develop invasive disease such as meningitis or bacteremia (Buttery J, 2002, Moxon ER, 1990b). The bacterium is classified as being encapsulated (typeable) or non-encapsulated (non-typeable, NT), a classification system which dates back to 1931. The typeable forms are divided into six types (a-f) depending on the structure of their capsule and can spread into the bloodstream to cause invasive disease.

## INTRODUCTION

The occurrence of typeable *H. influenzae* infections with severe outcome have declined dramatically since the type b *H. influenzae* conjugate vaccine was put into the routine vaccination protocol during the 1990's (Moxon ER, 2002, Weintraub A, 2003, Verez-Bencomo V, 2004).

This vaccine effectively handles the most common type b strain of *H. influenzae* but has no effect on the NT forms. NTHi are the second most common etiologic agent of acute OM and the most common cause of chronic otitis media. A number of surveys of healthy individuals indicate a rate of colonization between 40-80% among both children and adults (Rao VK, 1999). It has been shown that NTHi strains exhibit greater genetic diversity when compared to typeable *H. influenzae* strains. This diversity is visualized by species level-ribotyping (**Figure 2**), restriction fragment phoresis and multilocus sequence typing (MLST) (Cody AJ, 2003, Musser JM, 1986). Although most cases of NTHi easily can be treated with antibiotics such as ampicillin, resistance through development of  $\beta$ -lactamase is becoming increasingly common (Barry B, 1994). The physician visits and prescription of drugs to prevent NTHi infection now reach a staggering cost of one billion dollars per year in the US alone (Rao VK, 1999).



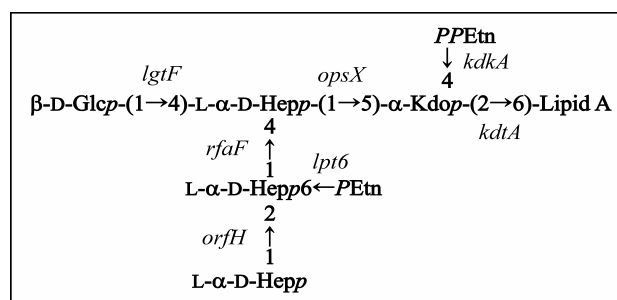
**Figure 2.** Species-level ribotyping dendrogram based on more than 400 *H. influenzae* strains. The dendrogram indicates the diversity seen in NTHi strains as compared to typeable strains.

## INTRODUCTION

---

### The fundamentals of *Haemophilus influenzae* LPS structure

LPS is categorized as being rough (R) or smooth (S) based on the absence or presence of an O-antigen polysaccharide (Alexander C, 2001, Hitchcock PJ, 1986). The work presented in this thesis deals with the (R) form of LPS as found in *H. influenzae*, thus without any repeating O-antigen structure. To visualize the LPS, it is convenient to separate the structure into segments. The lipid A region anchors the inner-core together with an inter-spaced 3-deoxy-D-manno-oct-2-ulsonic acid (Kdo) 4-phosphate residue to the outer membrane of the bacterium. The inner-core which comprises three L-glycero-D-manno-heptose (L,D-Hep) residues, referred to as Hep I (linked to Kdo), Hep II and Hep III (Schweda EK, 2007). Hep I is further substituted in the 4-position by a  $\beta$ -D-glucose commonly referred to as Glc I. The above mentioned elements are conserved among all *NTHi* strains studied to date and are therefore referred to as the conserved inner-core. Further conserved elements include a phosphoethanolamine (PEtn) residue in the O-6 position of Hep II and PPEtn at O-4 of Kdo (**Figure 3**). This template provides points for further oligosaccharide (OS) elongation from three distinct positions: Hep III (O-2 or O-3), Hep II (O-3) and Glc I (O-4/O-6). These OS extensions include various carbohydrate epitopes that form the heterogeneous outer-core of *H. influenzae*, a section which might vary both between and within strains. This heterogeneity is further enhanced by the inclusion of non-carbohydrate substituents including phosphate (P), phosphoethanolamine (PEtn), phosphocholine (PCho), acetate (Ac) and glycine (Gly) (see separate sections) (Schweda EK, 2007).



**Figure 3.** The conserved inner-core LPS structure of *H. influenzae*. Biosynthetic genes needed for the full assembly are indicated in italics (Hood DW, 2001, Tinnert AS, 2005, Wright JC, 2004).

## INTRODUCTION

---

This thesis deals with several *NTHi* strains of which two strains are particular important, namely strains 2019 and R2866 (INT1). *NTHi* 2019 was the first strain to have its major LPS element characterized in 1990 (Phillips NJ, 1992) and has since then been used in numerous labs across the globe, mostly to study biofilm formation (Greiner LL, 2004, Hong W, 2007, Jurecsek JA, 2005, Jurecsek JA, 2007). The full structural characterization of this strain is included in this thesis (**Paper III**).

For some time it was thought that the ability to pass into the blood stream was limited to typeable forms of *H. influenzae*, partly due to invasive non-typeable infections only being observed in children with defect immune system. Recently it has become apparent that *NTHi* can cause bacteremia and meningitis in otherwise healthy patients (Campos J, 2004, Cerquetti M, 2000, Cuthill SL, 1999, O'Neill JM, 2003). **Paper IV** contains the structural characterization of one of these invasive *NTHi* isolates, R2866, formerly known as INT1 (Nizet V, 1996).

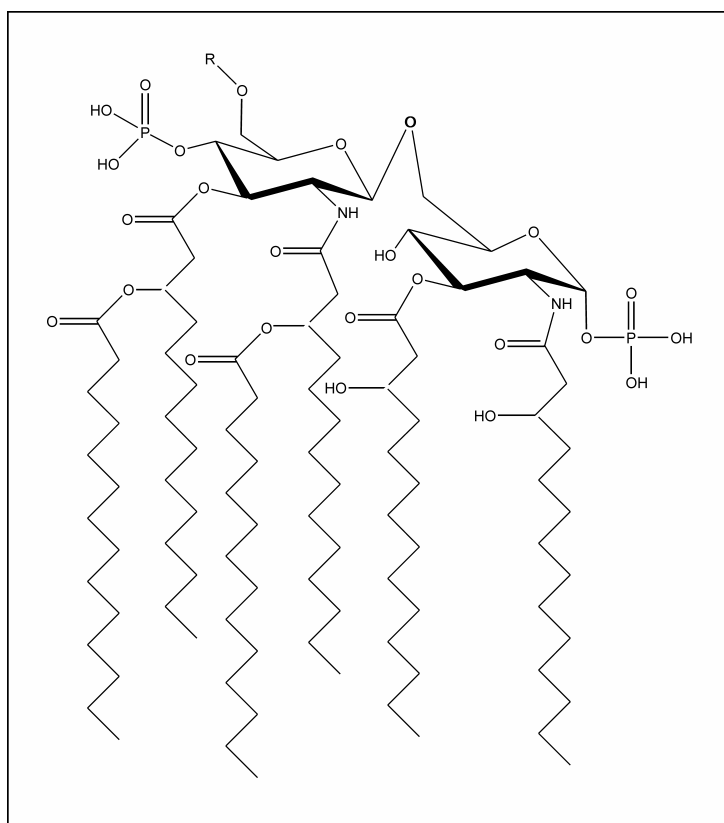
### Lipid A, structure and assembly

Over thirty years have passed since the lipid A portion of LPS was established as the endotoxically active segment of LPS. Since then, the structure of lipid A has been studied in several bacteria (Helander IM, 1988, Rietschel ET, 1994).

Lipid A serves as the hydrophobic anchor of LPS (Trent MS, 2004) with its composition as shown in **Figure 4**. The C2/C2' positions of the glucosamine backbone can be substituted by amide-linked 3-hydroxytetradecanoic acid chains while the C3/C3' are substituted through an ester-linkage. Furthermore, the fatty acids on C3' and C2' are esterified by tetradecanoic acid chains. Variation in structure is rare, and mostly related to the length of fatty acid chains (Helander IM, 1988, Mikhail I, 2005). The length of these fatty acid chains influence the toxic properties of lipid A (Rietschel ET, 1994). The most toxic lipid A structure identified was observed from *E coli* with C2 and C3 substituted by 3-hydroxytetradecanoic acid chains and C2' and C3' substituted by tetradecanoyloxydodecanoic acid, respectively. As of yet, two *NTHi* strains (1124 and 723) show differences related to *O*-acetylation in the 3-hydroxytetradecanoic acid group at C2 or C3 (Mikhail I, 2005), the function of *O*-acetylation of lipid A is still unclear.

## INTRODUCTION

Raetz and co-workers have carried out extensive studies of the enzymology and molecular genetics of lipid A synthesis using *E. coli* as a model system. They concluded that nine enzymatic steps are needed to assemble a full lipid A moiety, a pathway which is conserved among Gram-negative bacteria (Trent MS, 2004).



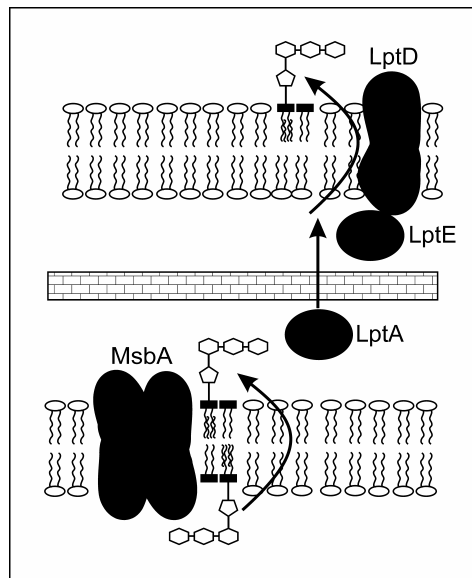
**Figure 4.** The structure of the lipid A part of the LPS molecule.



## INTRODUCTION

### Transport of LPS to the outer membrane

Prior studies have shown that LPS is made at the inner surface of the cytoplasmic membrane where the oligosaccharide is synthesized from the lipid A acceptor to form rough LPS (Ruiz N, 2009). Once the LPS has been synthesized, it needs to be transported to the outer membrane where lipid A anchors the LPS (Trent MS, 2004) to the outer membrane. The mechanisms through which components reach their final destination is largely unknown, though some progress has been made during the last years. The current understanding of transportation can briefly be summarized as follows: after being synthesized, LPS are flipped across the inner membrane by a flippase, *msbA*, LPS is then extracted from the inner membrane by a series of ABC transporters. One of these transporters, *lptA* functions as a chaperone which transports LPS to the assembly site on the outer membrane where *lptE* and *lptD* flip the LPS across the membrane (Ruiz N, 2009) (**Figure 5**).



**Figure 5.** LPS transport across the cell envelope. The figure indicates the flippases (*msbA*, *lptD* and *lptE*) and chaperones (*lptA*) needed to facilitate transport across the cell envelope to the final destination at the cell surface. Adapted from Ruiz N. *et al*, 2009.

## INTRODUCTION

### The variable outer-core

The diversity of carbohydrate expression as well as non-carbohydrate substituents in *NTHi* naturally means that the biosynthetic genes should be numerous. Much of the genetic blueprint of *NTHi* LPS expression was determined by sequencing data from the genome strain Rd (Fleischmann R. D. *et al.*, 1995), and relating gene function to structural features present in the type b strains Eagen (RM153) and RM7004 (Hood D. W. *et al.*, 1996a, Hood D. W. *et al.*, 1996b). The genetic blueprint for synthesis was later confirmed with strain Rd.

Moreover, the release of genome sequences from R2866 (**Paper IV**) and R2846 (Lundstrom S. L. *et al.*, 2008a, Harrison A. *et al.*, 2005, Erwin A. L. *et al.*, 2005) have facilitated yet another chance to identify novel genes.

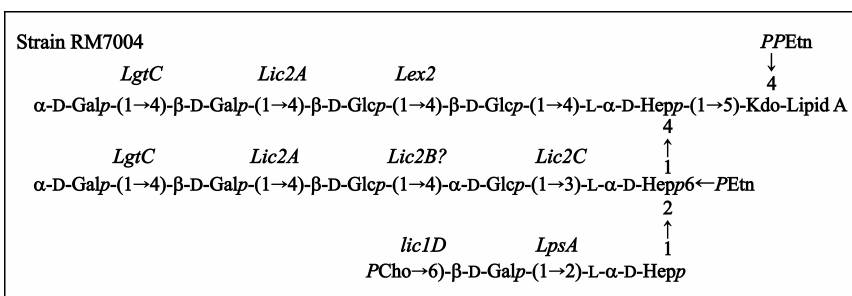
The diverse expression seen in outer core of *NTHi* provides a true challenge when one conducts structural elucidation of LPS. As there are three possible points of elongation from the conserved inner-core, the following section discuss each point separately with selected strains used to illustrate important gene functions and the corresponding epitopes.

The first possible point for OS extension is the conserved Glc I which can be substituted in O-4 or O-6 position by either a hexose or a heptose residues, respectively.

**O-4 substitution of Glc I** can give rise to the following epitopes: globotetraose [ $\beta$ -D-GalNAcp-(1 $\rightarrow$ 3)- $\alpha$ -D-Galp-(1 $\rightarrow$ 4)- $\beta$ -D-Galp-(1 $\rightarrow$ 4)- $\beta$ -D-Glcp-(1 $\rightarrow$ )] or [ $\beta$ -D-GalNAcp-(1 $\rightarrow$ 3)- $\alpha$ -D-Galp-(1 $\rightarrow$ 4)- $\beta$ -D-Galp-(1 $\rightarrow$ )] depending on which *lex2* transferase activity that is present in the strain (**Paper I/II**) (Schweda EK, 2007).

These epitopes can also be present in truncated versions, for instance globoside or lactose variants (see section on Phase variation). Furthermore, sialylated versions of these epitopes have been identified (see section on Sialic acid).

**Figure 6** shows an example of O-4 substitution on Glc I as seen in RM7004 (Masoud H, 2003). The required genes for full expression of the globoside epitope are indicated. The *lex2B* gene expresses a glucosyltransferase activity in this strain (Griffin R, 2003).



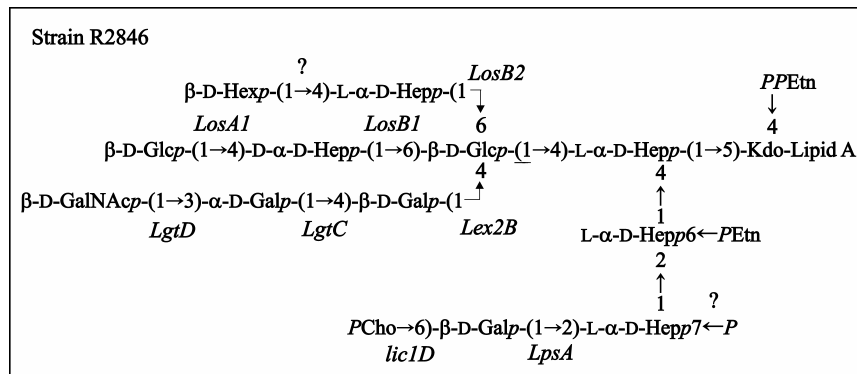
**Figure 6.** Structure of RM7004. Biosynthetic genes are indicated in italics.

## INTRODUCTION

The two new genome strains, NTHi R2866 and R2846 allowed for the characterization of **O-6 elongations from Glc I** and gave insights into the genes required for incorporation of heptoses (L,D-Hep or D,D-Hep) in the outer core. Genome homology comparison between *H. influenzae* and *H. ducreyi* identified four new genes related to LPS biosynthesis, two of which evidenced heptosyltransferases activity (*losB1* and *losB2*) in *H. ducreyi* (Post DM, 2007, Tullius MV, 2002).

Structural studies conducted in our lab indicated that R2846 produced both L,D and D,D-Hep in its outer core (Lundström SL, 2008a).

The genes responsible for the incorporation of these heptose residues were studied by construction of mutant strains R2846*losB1*, R2846*losB2* and R2846*losB1/losB2* and subsequent characterization of LPS. It was concluded that *losB1* controls the addition of D,D-Hep while *losB2* adds L,D-Hep (**Figure 7**). It has now been shown that some strains produce exclusively one of these heptosyltransferases (see **Paper IV**) while others can express both (Hood DW, unpublished data). The outer core heptose can be further substituted by hexose residues or a PCho residue (**Paper IV**) like [ $\beta$ -D-Glcp-(1→4), **Figure 7**], [ $\beta$ -D-Galp-(1→4)] and [ $\beta$ -D-GalNAcp-(1→3)- $\alpha$ -D-Galp-(1→4)- $\beta$ -D-Galp-(1→6)] (Lundström SL, 2008a, Månsson M, 2003a, Rahman MM, 1999). Moreover, sequence comparisons indicated that two genes (*losA1* and *losA2*), that are adjacent to the *losB* genes, encode predicted glycosyltransferases that are homologues of enzymes that add glucose units onto heptoses in *H. ducreyi* (Tullius M. V. et al., 2002). One of these transferase genes, *losA1* have been studied in NTHi R2846, and was evidenced to control the addition of glucose in O-4 position onto D,D-Hep (Deadman M.E. et al, unpublished results) (**Figure 7**).

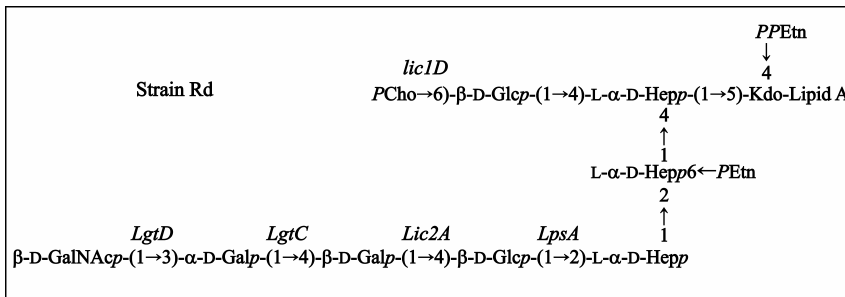


**Figure 7.** Structure of R2846 as identified by SL. Lundström and coworkers (Lundström SL, 2008a). Question marks indicate unknown genes. In this strain, *lex2B* expresses a galactosyltransferase.

## INTRODUCTION

To date, the following elongations have been identified for **O-3 substitution of Hep II**: [ $\alpha$ -D-Galp-(1→4)- $\beta$ -D-Galp-(1→4)- $\beta$ -D-Glcp-(1→4)- $\alpha$ -D-Glcp(1→3)] (Masoud H, 2003) and [ $\beta$ -D-GalNAcp-(1→3)- $\alpha$ -D-Galp-(1→4)- $\beta$ -D-Galp-(1→4)- $\alpha$ -D-Glcp(1→3)] (Lundström SL, 2007) or truncated versions thereof. **Figure 6** illustrates the required genes. The function of *lic2B* is still not fully understood, but it is possible that it expresses either a glucosyl or galactosyltransferase as described for *lex2B*.

**Hep III** has been shown to be substituted in either **O-2 or O-3** position by glucose or galactose depending on the sequence of the *lpsA* gene (Deadman ME, 2006a). To date, the following epitopes have been identified: [ $\beta$ -D-GalNAcp(1→3)- $\alpha$ -D-Galp-(1→4)- $\beta$ -D-Galp-(1→4)- $\beta$ -D-Glcp-(1→2)], [ $\beta$ -D-Galp-(1→2/3)] (Landerholm MK, 2004, Schweda EK, 2002) and [ $\beta$ -D-Galp-(1→4)- $\beta$ -D-Glcp-(1→3)] as well as truncated versions thereof (Månsson M, 2001, Schweda EK, 2003). This is illustrated by strain Rd shown in **Figure 8**.



**Figure 8.** Strain Rd with expression of a globotetraose epitope from Hep III.

It has been shown that *H. influenzae* also can synthesize a tetrasaccharide epitope which is added *en bloc*, not stepwise to Glc I. This mechanism is controlled by a cluster of genes in a region referred to as the high-molecular-weight glycoform (*hmg*) region. This region contains numerous genes that share homology with genes associated with O-antigen or polysaccharide synthesis in other Gram-negative organisms. To date, the following two epitopes have been identified: [Neu5Ac-(2→3)- $\beta$ -D-Galp-(1→4)- $\beta$ -D-GlcNAcp-(1→3)- $\beta$ -D-Galp-(1→4)] and related structure [PEtn→6)- $\alpha$ -D-GalNAcp-(1→6)- $\beta$ -D-Galp-(1→4)- $\beta$ -D-GlcNAcp-(1→3)- $\beta$ -D-Galp-(1→4)] (Hood DW, 2004).

## INTRODUCTION

---

### Host mimicry and phase variation

A ingenious feature of *NTHi* is host mimicry, where structural elements presented in the LPS mimic bio-structures found naturally within the host (Mandrell RE, 1992, Swords WE, 2000, Weiser JN, 1998b, Virji M, 1990). Host mimicry can be compared to a chameleon changing its color in nature making predators unable to recognize the food source since it mimics its natural environment. This contributes to the bacterium's ability to survive by preventing irradiation caused by the immune response. Host mimicry affects the bacterium's ability to adhere, persist and invade within the unfriendly host environment. These elements include carbohydrate structures like lactose [ $\beta$ -D-Galp-(1→4)- $\beta$ -D-Glcp-(1→)], digalactoside [ $\alpha$ -D-Galp-(1→4)- $\beta$ -D-Galp-(1→)], globotriose [ $\alpha$ -D-Galp-(1→4)- $\beta$ -D-Galp-(1→4)- $\beta$ -D-Glcp-(1→)] and globotetraose [ $\beta$ -D-GalNAcp-(1→3)- $\alpha$ -D-Galp-(1→4)- $\beta$ -D-Galp-(1→4)- $\beta$ -D-Glcp-(1→)] but also the incorporation of *PCho* and *N*-acetylneuraminic acid (Neu5Ac) in the LPS (see separate sections).

The intra-strain variation of the components in the outer-core region is in some part due to phase variation (Weiser J. N. *et al.*, 1990). This feature can be viewed as a defensive strategy taken by the bacteria enabling it to switch genes on/off reversibly and thereby produce a population of bacteria that are able to survive changing conditions. The mechanism is called slipped-strand mispairing and involves short tandem repeat sequences in the DNA double helix. Briefly, local denaturation and displacement of the strands followed by mispairing of complementary bases at the site of the short tandem repeat occurs. Several of these tandem repeats have been identified in *H. influenzae*, seven of which affect expression of sugar units or non-carbohydrate substituents. Seven phase-variable genes have been identified: *lic1*, *lic2A*, *lic3A*, *lic3B*, *lgtC*, *lex2* and *oafA*, all of which influence the virulence of the strain by inclusion of various substituents (Schweda EK, 2007).

In the following section, the phase variable genes *lic2A*, *lgtC*, *lex2* and the candidate gene *lic2B* will be discussed in detail, while genes related to sialic acid addition (*lic3A*, *lic3B*, *siaA* and *lsgB*), *PCho* addition (*lic1*) and acetate addition (*oafA*) will be discussed separately (see sialic acid and non-carbohydrate residues).

Phase-variable expressions related to epitopes identified from Glc I start by substitution of either a  $\beta$ -D-Glcp(1→4) or a  $\beta$ -D-Galp(1→4) residue depending on one key amino acid in the phase-variable *lex2* gene loci (**Paper I/II**). The subsequent sugar unit added depends on two other of the phase-variable genes, *lic2A* and *lgtC*. If *lex2* expresses a galactosyl transferase, then *lgtC* has the possibility to add  $\alpha$ -D-Galp-(1→4), thus forming a globotriose epitope at Hep I. If the opposite occur, that is that *lex2* encodes a glucosyl transferase, *lic2A*, in conjunction with *lgtC* add  $\alpha$ -D-Galp-(1→4)- $\beta$ -D-Galp-(1→4) onto the glucose.

## INTRODUCTION

---

Hep II elongations start off with  $\alpha$ -D-Glcp-(1→3), added by the *lic2C* gene. Further substitutions can occur, either with Gal or Glc (Lundström SL, 2007, Masoud H, 2003). This might involve the gene *lic2B* having a dual transferase activity as described for *lex2B* (unpublished data). Further elongations depend on *lgtC/lic2A* as discussed above and thus include digalactoside.

The situation at Hep III is more straight forward, with the first hexose residue (Glc or Gal) being added by the *lpsA* gene. Further extension is only possible if the hexose is a Glc (Deadman ME, 2006b). Possible extensions include the digalactoside epitope as described in both cases above.

The expression of digalactoside epitopes relate to virulence behavior (Weiser JN, 1998a). Strains which express two of these epitopes at once are more resistant to killing by naturally acquired antibodies and complement present in human serum which may allow the bacterium to persist within the host and alter its expression if microenvironment changes (Griffin R, 2005, Weiser JN, 1998a).

### **Sialic acid and implications for virulence**

Neu5Ac is an important structural element that influences the ability of the bacterium to avoid innate host immune responses (Schweda EK, 2007). All strains investigated to date express Neu5Ac in different quantities (Bauer SH, 2001). In each case, Neu5Ac functions as an anti-recognition molecule allowing the sialylated bacterium to hide from host immune defenses that otherwise rapidly would clear its unsialylated counterpart (Hood DW, 2001, Hood DW, 1999, Vimr E, 2002). The role of Neu5Ac *in vivo* was studied by Bouchet and co-workers (Bouchet V, 2003) using the well described chinchilla model of acute OM (Giebink GS, 1981, Karasic RB, 1985). It was concluded that Neu5Ac is a critical virulence factor in the pathogenesis of experimental OM in the chinchilla. These studies also demonstrated that sialylated glycoforms are needed predominantly during the early stages of OM onset, as the sialylated epitopes were lost after 9 days of infection (Lundström SL, 2008b). The production of sialylated LPS depends on the ability of the bacterium to scavenge Neu5Ac from the host (Severi E, 2005, Vimr E, 2000).

Recent studies implicate Neu5Ac as a critical component in biofilm formation, a feature of *NTHi* which now generates considerable amounts of publications (Armbruster CE, 2009, Figueira MA, 2007, Jurcisek J, 2005, Moriyama S, 2009). The experimental otitis media chinchilla model and scanning electron microscopy is used for detection of biofilm formation in these recent papers.

## INTRODUCTION

---

The microscopy technique allows one to visually look for biofilm formation in middle ear mucosa taken from the animal. Swords W.E. and co-workers used this approach to study the differences in biofilm formation between wild-type *NTHi* 2019 and mutant strain 2019*siaB* (Swords WE, 2004). The *siaB* gene encodes the CMP-Neu5Ac synthetase, thus mutations in this gene results in strains not capable of incorporating Neu5Ac in their LPS. The presence of biofilm formation is still under debate as one needs to characterize its presence by other techniques than electron microscopy (Moxon ER, 2008).

The biosynthetic genes responsible for the addition of Neu5Ac have been studied in detail (Fox KL, 2006, Månsson M, 2001). One of these genes, *lic3A* allows for the addition of  $\alpha$ -Neu5Ac-(2→3) onto lactose epitopes (Hood DW, 2001). Moreover, the close homolog *lic3B* is bifunctional, having the potential to add both [ $\alpha$ -Neu5Ac-(2→8)- $\alpha$ -Neu5Ac-(2→3)] and [ $\alpha$ -Neu5Ac-(2→3)] onto lactose epitopes found in the outer-core (Fox KL, 2006, Lundström SL, 2007, Yildirim HH, 2005a). To date, the following sialylated epitopes have been identified as controlled by *lic3A/lic3B*: [ $\alpha$ -Neu5Ac-(2→3)- $\beta$ -D-Galp-(1→4)- $\beta$ -D-Glcp-(1→)] and  $\alpha$ -Neu5Ac-(2→8)- $\alpha$ -Neu5Ac-(2→3)- $\beta$ -D-Galp-(1→4)- $\beta$ -D-Glcp-(1→).

Moreover, *NTHi* may express sialyllacto-*N*-neotetraose [ $\alpha$ -Neu5Ac-(2→3)- $\beta$ -D-Galp-(1→4)- $\beta$ -D-GlcNAcp-(1→3)- $\beta$ -D-Galp-(1→4)- $\beta$ -D-Glcp-(1→)] which is a result of *en bloc* addition and is influenced by two additional sialyltransferase genes, *lsgB* and *siaA* (Hood DW, 2004, Jones PA, 2002). Houliston and co-workers recently identified a novel epitope [ $\alpha$ -Neu5Ac-(2→8)- $\alpha$ -Neu5Ac-(2→3)- $\beta$ -D-Galp-(1→4)-L- $\alpha$ -D-Hep-(1→)] which substituted Glc I in O-6 position (Houliston RS, 2007). The incorporation of Neu5Ac in this epitope was controlled by *lic3B*.

### Non-carbohydrate substituents

The structural complexity of LPS expressed by *H. influenzae* is to some degree due to non-stoichiometric addition of non-carbohydrate substituents including *P*, *PEtn*, *PCho*, *Ac* and *Gly* (Schweda EK, 2007).

***PEtn* and *PPEtn*** have been found positioned at O-6 and O-4 on Hep II and Kdo, respectively. Moreover, a small subset of *NTHi* strains express an additional *PEtn* residue, either in O-3 position of Hep III (Månsson M, 2003b, Yildirim HH, 2005a) (**Paper III**) or in O-6 position on  $\alpha$ -GalNAc in the high-molecular glycoform expressed from Hep I (Houliston RS, 2007). Two genes related to these expressions have now been identified, *lpt6* being responsible for the phosphorylation of Hep II (Tinnert AS, 2005, Wright JC, 2004) and *lpt3* controlling addition of *PEtn* to Hep III (**Paper III**).

**Phosphate** substitution at O-4 or O-7 position of Hep III has also been demonstrated in a small number of strains (Lundström SL, 2008a, Masoud H, 1997), their relevant gene(s) and functionality are unknown.

## INTRODUCTION

---

**PCho** is commonly expressed by *H. influenzae* and can substitute several different positions. To date the following positions have been found: O-6 of Glc I, O-6 of glucose or galactose at Hep III, O-6 on  $\alpha$ -glucose at Hep II (Schweda EK, 2007) and O-7/O-6 on outer core heptoses (Lundström SL, 2008a) (**Paper IV**) (Fox KL, 2007, Houlston RS, 2007). Strains expressing two PCho moieties at once have also been identified (Fox KL, 2007).

As evidenced by the study performed by Fox and co-workers, the phase variable *licI* operon, which contain the *licIA-D* genes, control all aspects of PCho addition to *H. influenzae* LPS. NTHi strains 1158 (Houlston RS, 2007), R2846 (Lundström SL, 2008a), 11 and 16 (Landerholm MK, 2004) can express two PCho residues at once and have two distinct copies of *licID* allowing for two different positions To facilitate the incorporation of PCho, *licIB* transports choline from the host environment which then is phosphorylated by *licIA*. An activation takes place due to the pyrophosphorylase *licIC*. The now activated residue is transferred to its selected location by *licID*.

PCho plays a role in persistence of the bacterium on the mucosal surface of the nasopharynx, in part by mediating bacterial adherence to, and invasion of, the host epithelia (Weiser JN, 1998b). This function occurs through the receptor for platelet-activating factor (PAF) present on bronchial epithelial cells (Swords WE, 2000).

Furthermore, PCho expression has been associated with increased resistance to host antimicrobial peptide killing (Lysenko ES, 2000). The contrast to the benefit of PCho expression is that it also serves as a target for the serum component C-reactive protein (CRP), which once bound to PCho activates complement as described by Weiser and co-workers. Isogenic strains with PCho located at alternative positions in the LPS molecule shows different sensitivity to CRP, for instance clearance of the bacterium is faster when PCho substitutes the hexose at Hep III than if it substitutes Glc I.

**Acetates** have been found linked to O-3 or O-2 positions on Hep III or on the hexose adjacent to Hep III in O-6 or O-4 position (Cox AD, 2001, Landerholm MK, 2004, Månsson M, 2002, Risberg A, 1999). Moreover, O-2 of Hep I, O-3, O-4 or O-6 of Glc I can be substituted by acetates (Månsson M, 2003b, Schweda EK, 2000, Yildirim HH, 2005b). More recently, another acetylation site was discovered in the high-molecular glycoform of NTHi R2846 where GlcNAc was evidenced to be acetylated (Lundström SL, 2008a). The biological impact of O-acetylation has been studied by Fox and co-workers; they concluded that mutations in the phase variable gene *oafA* that directs addition of Ac to Hep III resulted in increased killing by normal human serum (Fox KL, 2005).

**Glycine** has been found linked to Hep III, Hep II, Hep I and Kdo in certain strains (Li J, 2001), the gene(s) responsible for the glycation are yet unknown as well as the benefit of such an expression.



## INTRODUCTION

---

### Colonization and infection

Infections caused by *H. influenzae* are complex processes which involve several different surface structures and a multitude of pathways. No single feature accounts for all disease-associated strains but rather a combination of factors which helps the bacteria to adhere, colonize and possibly spread to other parts of the host (Erwin AL, 2007).

The interaction between *NTHi* and the host begins with adherence of the bacteria to mucus or cells in the upper respiratory tract, in particular to the nasopharynx. The adherence to mucus is influenced by outer membrane proteins (OMP-2 and OMP-5) that bind mucin specific receptors (Reddy MS, 1996). Preferably, adherence is possible when the cilia have been damaged. This ciliostasis occurs when LPS molecules are released from the bacteria; these cause toxic effects mediated by lipid A (Rietschel ET, 1994, St Geme JW, 3rd, 2000). Furthermore, Janson and colleagues have shown that protein D induce ciliary damage (Janson H, 1999). At the next step the bacteria adhere to damage epithelial cells, several different adhesins are known to help *NTHi* perform this feat: some strains may produce pili, hair-like structures which enable adherence while others use high molecular weight adhesins HMW1 and HMW2 encoded by the genes *hmw1* and *hmw2* loci (Erwin AL, 2007). Approximately 75% of all *NTHi* strains contain these *hmw* genes (Rao VK, 1999). The remaining percentage of *NTHi* strains can use other adhesins such as IgD-binding outer membrane proteins Hia and Hsf as shown by Barenkamp *et al* or the autotransporter protein Hap (Barenkamp SJ, 1994, Barenkamp SJ, 1996, St Geme JW, 3rd, 2002).

The next challenge is to persist in the host which requires both the steady supply of nutrients and the ability to avoid host immune systems. The predominant immunoglobulin produced by mucosal tissue is IgA which inhibits adherence and invasion (Kilian M, 1979, St Geme JW, 3rd, 2000). *H. influenzae* does so by production of an extra cellular endopeptidase called IgA1 protease which through its complex action prevents the action of IgA. It has been shown that approximately 97% of all strains produce this endopeptidase (Poulsen K, 1992). Moreover, micro colony formation increases resistance to bacteriostatic compounds such as lysozyme, peroxidases and lactoferrin (St Geme JW, 3rd, 2000). These colonies also block access of antibodies to individual organisms, thus impeding antibody dependent killing and phagocytosis. OMP-2, OMP-5, IgA1 protease and pili also undergo irreversible antigenic drift which further enables the bacterium to avoid the immune response (St Geme JW, 3rd, 2000).

The endotoxin response from the host triggers when bacteria die and the LPS is set free from the outer membrane. Lipid A is the endotoxically active portion of the LPS, its structure influences the toxicity of the bacterium (see lipid A section). The innate and

## **INTRODUCTION**

---

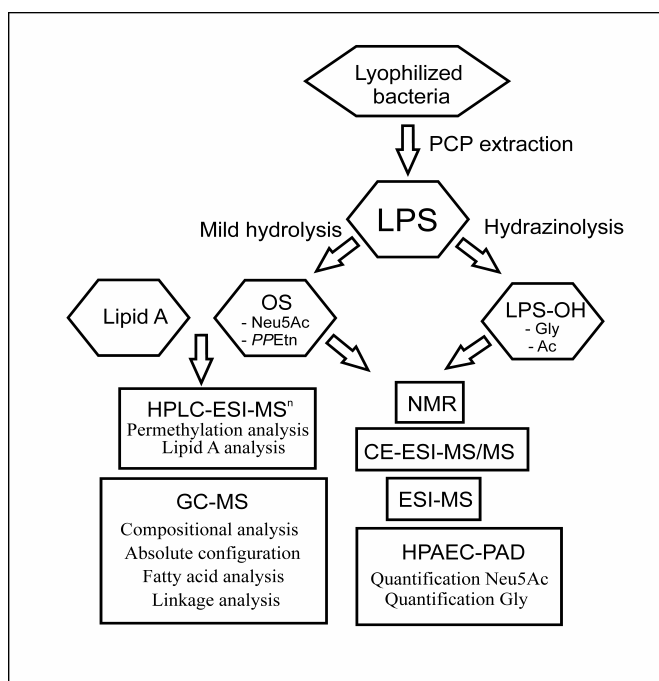
adaptive immune responses are activated once the endotoxin is set free (Alexander C, 2001, Holst O, 1996) which in moderate infections lead to mild fever and irradiation of the bacteria. More severe consequences such as shock, sepsis and tissue death can occur if high amounts endotoxin are released.

Invasive disease, such as meningitis occurs when the successfully colonized bacteria penetrate into the bloodstream through the tight junctions in nasopharyngeal epithelial cells (Farley MM, 1986). The next step involves a process known as transcytosis where the bacteria cross the blood-brain barrier, ultimately causing meningitis (Virji M, 1992).

## MATERIALS AND METHODS

To be able to study both the genetics underlying LPS expression as well as elucidate the structures of various *NTHi* strains one has a toolbox with both microbiology techniques and chemical/analytical methods at hand. Several daunting tasks need to be solved, including growing strains, constructing mutant strains (see bacterial cultivation and construction of mutant/transformant strains) and of course putting the structural puzzle together. To conduct a full structural characterization one needs to consider which monosaccharides are present in the sample, the size of their ring, how they are linked to each other as well as the absolute and anomeric configuration. Since *H. influenzae* also express non-carbohydrate residues, their position and linkage should also need to be considered. Moreover, possible branching patterns and the sequence of residues in the various glycoforms need to be resolved.

LPS is extracted and isolated from lyophilized bacteria. To solve solubility issues related to working with native LPS and to simplify elucidation one can subject LPS to either *O*-deacylation or mild hydrolysis to obtain *O*-deacylated LPS (LPS-OH), oligosaccharide (OS) material and liberated lipid A, respectively. These materials can be either chemically degraded, derivatized or directly analyzed by various techniques such as mass spectrometry (MS) or nuclear magnetic resonance (NMR) (Figure 9). These reactions and methods will be discussed in detail in the following section.



**Figure 9.** Methods and instruments used for structural characterization of LPS.

## MATERIALS AND METHODS

---

### Bacterial cultivation and construction of mutant/transformant strains

In order for *H. influenzae* to grow two important nutrients are needed, namely haemin ( $10 \mu\text{g}\cdot\text{mL}^{-1}$ ) and NAD ( $2 \mu\text{g}\cdot\text{mL}^{-1}$ ). Normal growth conditions include brain heart infusion broth (BHI) which is supplemented by these growth factors and cultured at  $37^\circ\text{C}$ . The bacteria are allowed to grow until late logarithmic phase after which they are killed and lyophilized.

Several mutant strains were used in this thesis. In these strains various genes were inactivated in order to study their potential role in LPS expression or to simplify structural elucidation. The strains were mutated by transformation with linearised plasmid containing the gene in question interrupted by the insertion of a kanamycin resistance cassette. For instance, if the *lex2B* gene is knocked out in strain 2019, the strain is then named 2019*lex2B*. PCR and Southern blot analyses are performed to ensure that the mutant strains are functioning as intended (Herriott RM, 1970, Hood DW, 1996a). In **Paper II**, transformant strains were constructed by inserting the *lex2* locus from one *NTHi* strain into another. Briefly, the *lex2* loci from *NTHi* strain 1008 expressing a glucose linked to Glc I naturally was inserted into *NTHi* strain 1124 which normally expresses a galactose in the same position. This transformant strain is named 1124*lex2Glc*<sub>1008</sub>, thus implying that this transformant strain now should express glucose instead of galactose.

### Extraction of LPS

The phenol:chloroform:light petroleum (PCP, 2:5:8) method is used in order to extract the LPS from the surface of the outer membrane (Galanos C, 1969). This method produces good quantities of pure material from (R)-form LPS.

Briefly, the lyophilized bacteria are stirred in the PCP solution ( $21^\circ\text{C}$ , 24 h) and then subjected to centrifugation. The supernatant is stored and the pellet is redissolved in fresh PCP solution, stirred and centrifuged using the same conditions as mentioned previously. The supernatants are combined and reduced by rotary evaporation until only the phenol phase remains. A mixture of acetone/diethyl ether (5:1) is added which forces the crude LPS to precipitate. LPS is collected by centrifugation and washed three times with acetone in order to remove all traces of phenol, since failure to do so might lead to unintended cleavages of the complex (Masoud H, 1997). The LPS still contain some impurities, for instance RNA, which is effectively removed by ultracentrifugation ( $75\ 000\ \text{g}$ ,  $4^\circ\text{C}$ , 16 h) and then lyophilized.

## MATERIALS AND METHODS

---

### Preparation of material

Problems related to solubility of native LPS is solved by either *O*-deacylation (hydrazinolysis) which removes ester-linked fatty acids from the lipid A or delipidation (mild acid hydrolysis) which removes lipid A. These two reactions yield two distinctly different materials, *O*-deacylation affords LPS-OH while delipidation gives insoluble lipid A and OS material.

### Hydrazinolysis of LPS

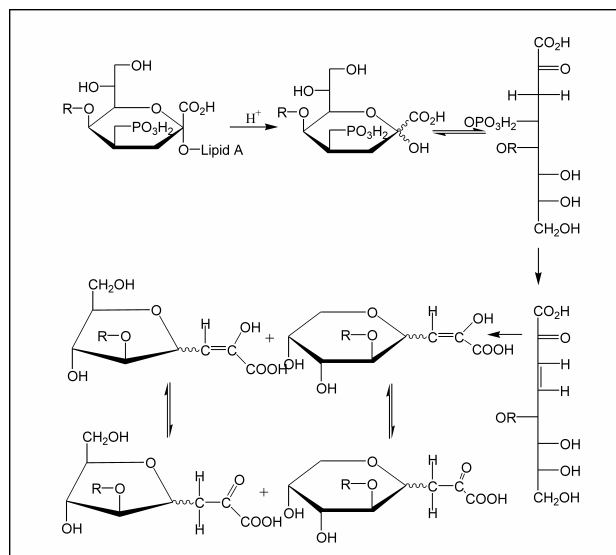
Mild *O*-deacylation of LPS with anhydrous hydrazine results in LPS-OH material with losses of ester-linked glycine and acetate groups as well as the ester-linked fatty acid chains substituting lipid A (Holst O, 1991). The loss of fatty acid chains increase the solubility. LPS-OH still contains Neu5Ac which is primary detected by capillary electrophoresis electrospray ionization mass spectrometry (CE-ESI-MS). Briefly, LPS is treated with hydrazine (40 °C, 1 h) after which the reaction mixtures is cooled on ice and remaining hydrazine is neutralized by the addition of cold acetone. Precipitated LPS-OH is washed with acetone by centrifugation.

### Hydrolysis of LPS and preparation of lipid A

Following mild acid hydrolysis of LPS with acetic acid (pH 3.1, 100 °C, 2 h), under simultaneous reduction with borane-*N*-methylmorpholine complex, yields both liberated lipid A and reduced core OS. These two components are isolated from each other by centrifugation, the OS material is further purified and fractionated by gel-permeation chromatography (GPC, Bio-Gel: P-4, 800-4000 Da and G-15 ≤ 1500 Da). The *PPE*tn residue at O-4 position of Kdo is lost during this treatment, this loss occurs through β-elimination and result in the formation of several anhydro-Kdo (*AnKdo*-ol) forms (**Figure 10**) (Auzanneau F-I, 1991, Danan A, 1982, Schweda EK, 1993). In addition, Neu5Ac residues are lost due to the acid labile nature of their ketosidic linkage.

Lipid A is purified by partition between chloroform:methanol:water (2:1:1) (Mikhail I, 2005). Lipid A is subjected to electrospray tandem mass spectrometry (ESI-MS<sup>n</sup>). The fatty acids are analyzed by converting them to their corresponding fatty acid methyl esters (FAMES) (1: HCl, 100 °C, 4 h, 2: MeOH, acetyl chloride, 80 °C, 16 h) and detected by gas chromatography – mass spectrometry (GC-MS).

## MATERIALS AND METHODS



**Figure 10.** The formation of several anhydro forms of Kdo (*AnKdo-ol*) upon mild acid hydrolysis (Auzanneau FI, 1991, Danan A, 1982, Schweda EK, 1993).

### Dephosphorylation and deacylation of OS

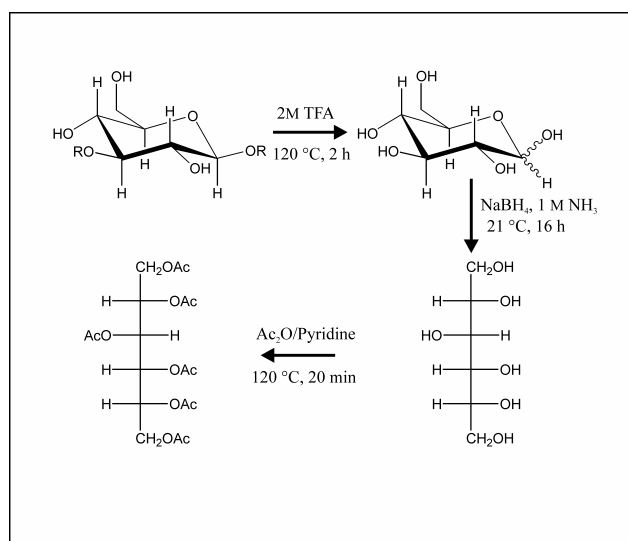
Dephosphorylation of OS material (48 % HF, 4 °C, 48 h) is used in order to remove all phospho-related substituents. It is of extreme importance to keep the temperature under control, as higher temperatures can result in cleavages of glycosidic linkages elsewhere in the LPS. Deacylation can be performed under mild conditions (1 M NH<sub>3</sub>, 21 °C, 16 h) to remove acetate and glycine groups. Comparing native OS with its deacylated counterpart is sometimes the only way to confirm the presence of Ac and Gly groups. This may also enable the detection of trace amounts of compositions which were originally overshadowed by the presence of non-carbohydrate substituents in the native material.

### Compositional and absolute configuration analysis

The identity of the monosaccharides can be established by a three-step reaction known as sugar or compositional analysis. The first step is acid hydrolysis (2 M TFA, 120 °C, 2 h) to cleave the glycosidic bonds between the sugar units after which a reduction (NaBH<sub>4</sub>, 1 M NH<sub>3</sub>, 21 °C, 16 h) is performed which converts the monomers into their corresponding alditols. These alditols are finally acetylated (Ac<sub>2</sub>O/Pyridine, 1:1 v/v,

## MATERIALS AND METHODS

120 °C, 20 min) forming alditol acetates which have volatility suitable for analysis by GC-MS (Sawardeker JS, 1965) (**Figure 11**). The drawback with this analysis is that Neu5Ac and Kdo are not detected as they are degraded during the hydrolysis. Also, *N*-acetylhexoseamine (HexNAc) will be underrepresented due to partial *N*-deacetylation and differences in response factors as compared to hexoses.



**Figure 11.** Principle steps of compositional analysis. The alditol acetate obtained for glucose is illustrated in the figure.

Another important characteristic of carbohydrates is their absolute configuration. To be able to state in which form they occur, one needs a way to convert them from enantiomers to diastereomers (Gerwig GJ, 1979). This is done by the Fischer glycosidation with a chiral alcohol. OS samples are hydrolyzed (2 M TFA, 120 °C, 2 h) and subjected to the chiral alcohol ((+)-2-butanol, acetyl chloride, 85 °C, 8 h). After this follows an acetylation step (Ac<sub>2</sub>O/pyridine, 1:1, v/v, 120 °C, 20 min). If the sample contains *N*-acetylhexoseamine residues one needs to perform an extra *N*-acetylation step (Ac<sub>2</sub>O, 21 °C, 4 h).

## MATERIALS AND METHODS

---

### Methylation and permethylation analysis

The linkage positions can be established using the methylation procedure (Blakeney AB, 1985) which includes methylation of free hydroxyl groups. OS material is dissolved in dimethyl sulphoxide (DMSO, 21 °C, 24-48 h) and then subjected to a modified version of Hakomori's reactions (Hakomori S, 1964), with lithium methylsulphinyldimethylmethanide (BuLi, 40 °C, 1 h) and methyl iodide (MeI, 0 °C during addition, 21 °C, 16 h). The methylated components are recovered on a SepPak C18 cartridge and can now be used for permethylation analysis (**Figure 9**) provided they were dephosphorylated prior to permethylation. Dephosphorylated and permethylated OS samples are analyzed by high performance liquid chromatography – electrospray – multiple step tandem mass spectrometry (HPLC-ESI-MS<sup>n</sup>) to obtain crucial information about branching and sequence of the glycoforms present. In order to continue with linkage analysis, the sample is subjected to sugar analysis. The obtained derivatized material is referred to as partially methylated alditol acetates (PMAAs) and are analyzed by GC-MS. To be able to unambiguously identify the former position of the carbonyl group in the sugar of interest, the reduction step during the sugar analysis must be performed with NaBD<sub>4</sub>.

### Mass spectrometry

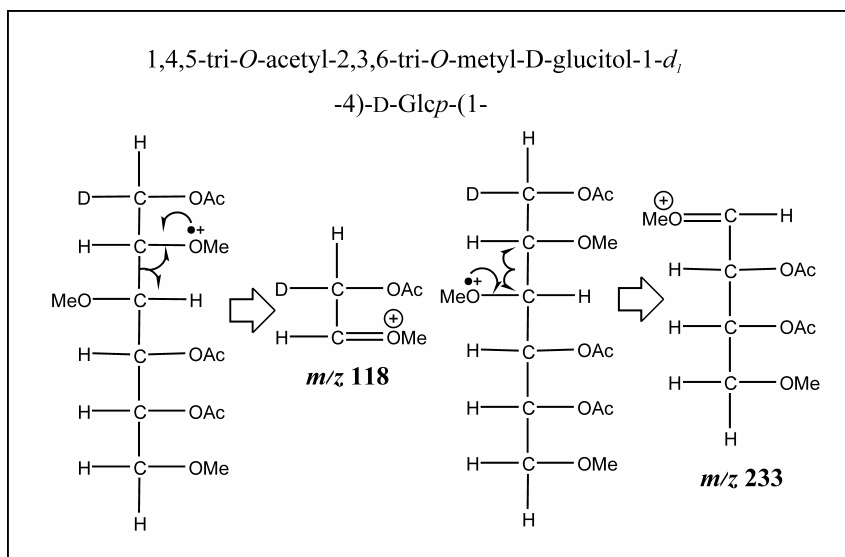
The founding father of mass spectrometry, J. J. Thomson performed his first simple experiments in 1912. Since then, mass spectrometry as a technique has undergone countless improvements. One of the milestones was the coupling of chromatography to mass spectrometry which not only provided means of separation for complex mixtures, but also quantification. It is fair to say that the field of structural characterization of LPS would be impossible without techniques involving mass spectrometry. This thesis briefly presents the techniques we use and what information that can be gained from them.

### GC-MS

Compositional, linkage, fatty acid and absolute configuration analysis are all performed on GC-MS, providing information about the identity of sugars and fatty acid chains as well as the saccharides linkage and absolute configuration (Gerwig GJ, 1979, Hellerqvist CG, 1990, Mikhail I, 2005). The alditol acetates, PMAAs and FAMEs are identified according to their retention times in the GC chromatogram, but also through their fingerprint mass fragmentation obtained as they are fragmented in the mass spectrometer fitted with an electron impact (EI). Each derivative has its own preferred fragmentation pathways depending on the positions of methyl and acetate groups, with the breakage between two methyl groups being the most common as observed in linkage analysis. **Figure 12** demonstrates the principle for interpretation of 4-substituted Glc. Furthermore, secondary fragmentation such as elimination of acetic acid and acetic acid anhydride may provide additional information.



## MATERIALS AND METHODS



**Figure 12.** Characteristic mass fragmentation patterns as observed for 4-substituted Glc. Indicated are key fragments  $m/z$  118 and 233.

### ESI-MS

The ESI-MS technique works by applying an electric field at atmospheric pressure to a liquid which passes through a capillary at low flow rate. The electric field is produced by applying a potential difference between the capillary and a counter electrode which causes a charge accumulation at the surface of the liquid. As the solvent evaporates, the charges close in on each other causing the liquid to break, forming a highly charged spray of droplets to which the ionization technique's name relates (Hoffmann DE, 1996).

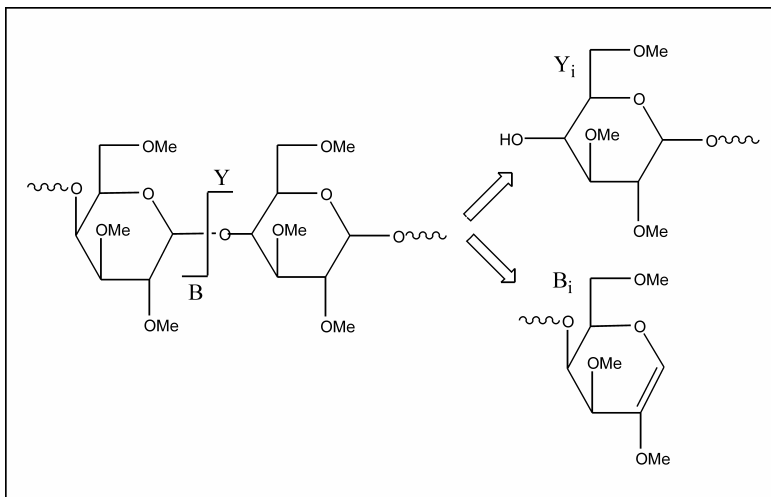
In this thesis, ESI-MS is used to determine the compositions of sugars and non-carbohydrate residue present in the LPS (Schweda EK, 2003). These compositions can be viewed as molecular ions with multiple charges depending on the amount of ionizable sites on the molecule. The ions are interpreted as  $[(M-nH)^n]$  or  $[(M+nH)^n]$  depending on whether the sample is run in negative or positive mode. Several solvents can be used, most commonly mixtures of acetonitrile, water and methanol with additions of  $NH_3$  (negative mode) or HOAc (positive mode) to ensure that proper ionization takes place. Some strains might express high amounts of non-carbohydrate residues (Gly/Ac) which clutters the spectra making interpretation hard. It is then possible to deacylate these samples in order to remove these residues.

## MATERIALS AND METHODS

### HPLC-ESI-MS<sup>n</sup> and CE-ESI-MS/MS

By including a quadrupole ion trap (QIT) one has the possibility to run tandem ESI-MS experiments (ESI-MS<sup>n</sup>) with collision induced dissociation (CID) which enables sequence and branching information to be established. The compounds are analyzed as pseudo molecular ions  $[M+Na]^+$  which are formed through the use of a solvent containing 1 mM NaOAc in methanol:water. Fragmentations by CID results primary in cleavage of the glycosidic bond with oxygen atom retention on the reducing-end fragment (Viseux N, 1998).

The nomenclature for these type of fragmentations were first suggested by Domon and Costello (Domon B, 1988), thus producing either an B<sub>i</sub> fragment with sodium ion retained on the nonreducing end or Y<sub>i</sub> fragment with retention on the reducing end (Figure 13).



**Figure 13.** CID fragmentation of glycosidic bonds as observed from permethylated OS. The figure indicates the formation of the Y<sub>i</sub> and B<sub>i</sub> fragments.

Branching and sequencing data from the dephosphorylated and permethylation OS samples are obtained through interpreting fragmentation patterns. **Table 1** summarizes the permethylated monosaccharide residue masses. This information is used to make decisions on which ions to fragment further through MS<sup>n</sup> experiments.

## MATERIALS AND METHODS

---

Monosaccharide	Mass ( <i>m/z</i> )
Terminal Hex	218
Monosubstituted Hex	204
Disubstituted Hex	190
Terminal Hep	262
Monosubstituted Hep	248
Disubstituted Hep	234
Terminal HexNAc	259
Monosubstituted HexNAc	245

**Table 1.** Permethylated monosaccharide residue masses.

CE-ESI-MS/MS can be used to determine the arrangements of Neu5Ac and non-carbohydrate residues in both native LPS-OH and OS samples (Li J, 2007). Presence of Neu5Ac is typically confirmed by precursor ion monitoring experiments scanning for the loss of Neu5Ac (*m/z* 290, negative mode) or Neu5Ac-H<sub>2</sub>O (*m/z* 274, positive mode).

### NMR spectroscopy

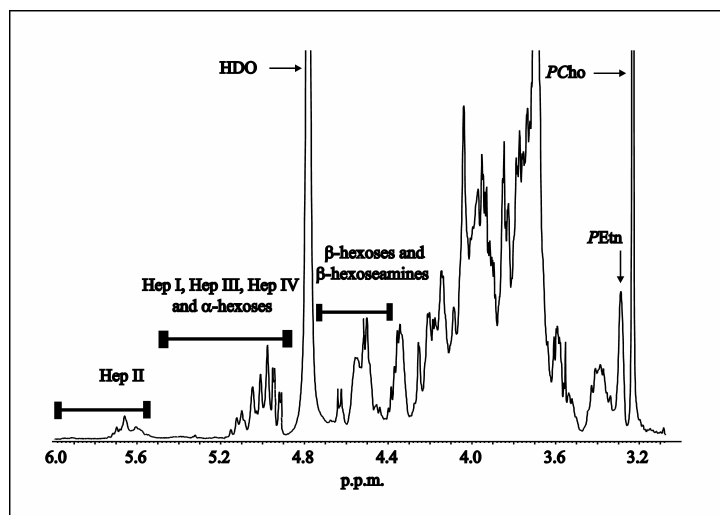
NMR spectroscopy plays a pivotal role in the characterizations of LPS. The technique provides extensive structural information such as sugar identity, absolute configuration, sequence/branching data and even positions of non-carbohydrate residues. In addition to all these great benefits, NMR is a non destructive technique. In **Paper II**, NMR analysis was crucial in order to assess the identity of the hexose residue linked to Glc I, and made it possible to conclude that transformations from Glc to Gal or vice versa were successful. The downside of NMR is twofold, the first major drawback is poor sensitivity. The second drawback is that complex samples often yield inheritable complicated 1D spectra with extensive overlap of signals. The development of various technical gadgets and improvements including cryo probes and stronger magnets have made these drawback less of a problem. The problems with overlapping signals can in part be resolved through the use of two-dimensional (2D) experiments where the signals are resolved in one more dimension. In this thesis, deacylated OS samples were dissolved in D<sub>2</sub>O and experiments were performed at 22-25 °C (HDO, δ<sub>H</sub> 4.8). Chemical shift data are reported in p.p.m. referenced to internal sodium 3-trimethylsilylpropanoate-*d*4 (δ<sub>H</sub> 0.00, <sup>1</sup>H NMR), external acetone (δ<sub>C</sub> 30.1, <sup>13</sup>C NMR) and external phosphoric acid (δ<sub>P</sub> 0.00, <sup>31</sup>P NMR).

## MATERIALS AND METHODS

### 1D $^1\text{H}$ NMR

OS samples analyzed in this thesis, as well as complex carbohydrate structures in general, typically give  $^1\text{H}$  signals in three distinct regions. The anomeric protons resonate in the anomeric region  $\delta$  4.4-6.0 with the conserved inner-core heptoses located between  $\delta$  5.0-6.0 (Cox AD, 2002, Masoud H, 2003, Masoud H, 1997, Månsson M, 2003a, Rahman MM, 1999, Schweda EK, 2002, Yildirim HH, 2005a). Hep II is generally found shifted downfield and may vary depending on the substitution patterns on Hep III (**Paper III**) (Schweda E. K. *et al.*, 2003). A distinct splitting is often observed for the H-1 signals occurring for Hep I and Hep II due to the micro heterogeneity of the *AnKdo-ol* moiety. Briefly, hexoses and heptoses in  $\alpha$ -configuration are observed at  $\delta$  4.9-5.5 while their counterpart, hexoses and hexoseamines in  $\beta$ -configuration are seen at  $\delta$  4.4-4.8. The most complex region in the NMR spectra is observed at  $\delta$  3.2-4.4, in this region all ring protons resonate which results in extensive overlap (**Figure 14**). Vicinal coupling constants ( $^3J_{\text{H-1,H-2}}$ ) provide information about the sugars anomeric conformation, briefly hexoses and hexoseamines in *gluco* or *galacto* configuration give vicinal coupling constants around 4 Hz when in  $\alpha$ -configuration and  $\sim 8$  Hz when in a  $\beta$ -configuration. Heptoses on the other hand side, being in  $\alpha$ -*manno* configuration give  $^3J_{\text{H-1,H-2}}$  around 2 Hz.

Non-carbohydrate residues substituting the LPS can also be identified due to specific sharp signals, for instance *PCho* produces a singlet at  $\delta$  3.23 arising from its methyl protons. The ethylene protons from *PEtn* also give rise to an intense signal at  $\delta$  3.29.

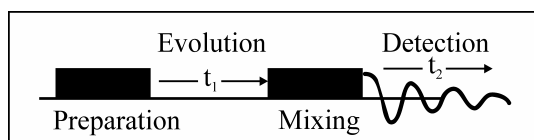


**Figure 14.** 1D NMR spectrum of *NTHi* strain R2866. Indicated are the specific shift regions for  $\alpha$ -heptoses and  $\alpha/\beta$ -hexoses and hexoseamines found in the outer-core. The signals arising from the methyl and ethylene protons of *PCho* and *PEtn* are also indicated.

## MATERIALS AND METHODS

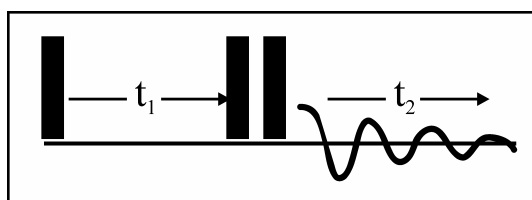
### 2D homonuclear $^1\text{H}$ - $^1\text{H}$ correlation spectroscopy

The introduction of 2D NMR in structural elucidation has simplified the daunting task of producing full characterizations of complex biomolecules in a profound way. One simple way to represent just about any 2D experiment setup is by dividing the experiment into four regions as shown in **Figure 15**. The preparation period could be as simple as a standard  $90^\circ$  pulse to generate transverse magnetization as used in a correlation spectroscopy (COSY) experiment. The evolution period is where coherences are allowed to evolve, without any observations being made. Next comes the mixing period where one aims to manipulate the existing coherences into an observable signal which is detected in the last period ( $t_2$ ). It is also possible to transfer magnetization from one spin to another, for instance from  $^1\text{H}$  to  $^{13}\text{C}$  (Keeler J., 2005).



**Figure 15.** General scheme for 2D NMR.

COSY experiments give diagonal peaks and cross peaks, the cross peaks is a result of scalar coupling. In principal this means that the anomeric proton of each sugar unit produces a diagonal multiplet with its H-2 proton visible as an off-diagonal cross multiplet, making it possible to trace out the J-coupling network in the molecule. When applying this theory to saccharide units, the overlapping signals from the ring protons often make it impossible to follow signals beyond H-2 (Duus J, 2000). The H-2 proton shifts obtained from COSY experiment can often help to determine if the spin system is related to a glucose or galactose, since the latter often has its H-2 at slightly higher chemical shift. Double-quantum filtered COSY (dqf-COSY) (**Figure 16**) is often used in this thesis. Briefly, this experiment differs from normal COSY pulse sequences by inclusion of one additional  $90^\circ$  pulse just before detection in  $t_2$ . The consensus of this is that only anti-phase peaks with absorptive line shapes are seen which results in a much better spectrum with balance between cross and diagonal peaks (Keeler J., 2005).



**Figure 16.** The pulse sequence for dqf-COSY.

## MATERIALS AND METHODS

---

Total correlation spectroscopy (TOCSY) experiments provide crucial information in this thesis, the basics relate to COSY but with the significant difference that cross peaks can be observed between spins which are connected by an unbroken chain of couplings. In practice this means that the anomeric proton is displayed as a diagonal peak, with several cross peaks observed for H-2 and the remaining protons (Keeler J., 2005). The key factor in setting up a well functioning TOCSY experiment is deciding on sufficient spin lock time (mixing period in general 2D experiments), in this thesis, 180 ms is commonly used. When running TOCSY experiments on saccharide chains, the dihedral angle between the protons relate to the  $^3J_{H,H}$  coupling constants which limits the transfer of magnetization. What this means in practice is that if the coupling constants are small, the magnetization cannot be transferred effectively to the next proton. Small coupling constants between H-4 and H-5 in a *galacto* configured sugar residue as well as between H-1 and H-2 of sugars in *manno* configuration prevent the transfer of magnetization. This typically makes determination of H-3→H-7 of heptoses and H-5/H-6 chemical shifts of Gal/GalNAc residues hard to determine through TOCSY experiments.

Nuclear Overhauser enhancement spectroscopy (NOESY) experiments are one of the most valuable experiments used in this thesis. Not only does it provide information on the anomeric configuration and ring form (pyranose/furanose), it also shows inter/intra-residue NOE connections between connecting sugar units allowing the sequence elucidation (**Table 2**). These experiments rely on protons being close to each other in space, typically between 3-5 Å. When two spins are within close proximity to each other, a feature called cross relaxation comes into play. In practice this means that these spins give rise to a cross peak, thus making it possible to trace the sequence of the included saccharide units. The magnetization is transferred through dipolar coupling during the mixing time. This value is extremely important to control, since too large values may give rise to unwanted effects such as spin diffusion. Spin diffusion can briefly be described as additional inter-residue NOE connections, protons which are not close to each other but still show cross peaks (Keeler J., 2005).

Anomeric proton	Intra-residue connection	Inter-residue connection
Hep I	H-2	
Hep II	H-2	H-3 Hep I, H-1/H-2 Hep III
Hep III	H-2	H-1/H-2 Hep II, H-1 Hex
Glc I	H-3, H-5	H-4/H-6 Hep I
$\beta$ -Hex (1→2/3)-Hep III	H-3, H-5	H-1/H-2/3 Hep III
$\alpha$ -Glc (1→3)-Hep II	H-2	H-3 Hep II
Hep IV (L,D or D,D)	H-2	H-6 <sup>ab</sup> Glc I

**Table 2.** Examples of inter/intra-residue connections observed in OS samples during a NOESY experiment. The hexose residue linked to Hep III can be a Glc or Gal residue linked in either O-2 or O-3 position.

## MATERIALS AND METHODS

---

### 2D heteronuclear $^1\text{H}$ - $^{13}\text{C}$ and $^1\text{H}$ - $^{31}\text{P}$ correlation spectroscopy

The heteronuclear correlation spectroscopy experiments are used to provide the  $^{13}\text{C}$  and  $^{31}\text{P}$  chemical shift data needed for full structural elucidation of LPS. Heteronuclear multiple quantum coherence (HMQC) correlates the proton shift data obtained from the homonuclear experiments with  $^{13}\text{C}$  or  $^{31}\text{P}$  shifts. A benefit of these experiments is that magnetization initially is on the  $^{13}\text{C}/^{31}\text{P}$  spin (S-spin) but later it is transferred to the  $^1\text{H}$  spin (I-spin) where detection takes place, yielding an increase signal detection (Keeler J., 2005).

The anomeric carbon region in an HMQC spectra is observed between  $\delta$  95-105 p.p.m., the actual value can help to determine anomeric configuration. Residues in  $\beta$ -configuration have higher C-1 shifts than the corresponding residue in  $\alpha$ -configuration. The ring carbons are found at  $\delta$  65-80, they can also differ depending on whether they are substituted (shifts downfield) or not, a crucial piece of information which is used in this thesis. Signals arising from the methyl carbons of *P*Cho, *O*-acetyl and *N*-acetyl groups are observed at  $\delta$  ~55, ~21 and ~23 p.p.m., respectively.

HMQC experiments with correlation between proton and phosphorus is used in order to determine the position of phosphor-related species like *P*Cho and *PEtn*. In these experiments, the  $^3\text{J}$  scalar couplings between the proton and phosphorus atom give rise to cross peaks, moreover couplings to the ethylene protons of these residues are visible (**Paper III/IV**).

### HPAEC-PAD and GPC

GPC is used in order to purify core OS material obtained after delipidation. Glycoforms of larger size eluate first as leading fractions, while smaller glycoforms eluate later. The samples are detected through their refractive index by using a refractometer. Furthermore, separation upon compositions containing *P*Cho can be achieved.

High-performance anion exchange chromatography (HPAEC) linked to pulsed amperometric detection (PAD) is used to quantitate the amount of Neu5Ac present in LPS-OH samples (Bauer SH, 2001). Neu5Ac is liberated from LPS-OH by enzymatic treatment (neuraminidase) or hydrolysis by HCl (0.1 M HCl, 80 °C, 1 h). Separation is achieved under basic conditions (pH ~12-13) by using a column containing pellicular anion exchange resin beads with increasing amounts of sodium acetate. The signal is measured by the PAD, this detector notes the electrical current generated by oxidation. By employing this method, it is possible to obtain excellent signal to noise ratios of quantities ranging from 1000-2 picomoles/25  $\mu\text{L}$  injected sample.

## **AIM OF THE THESIS**

---

The aim of the thesis was to elucidate the LPS structure of several disease associated *NTHi* strains and increase the knowledge concerning the biosynthetic genes needed for specific LPS expression patterns.

Specific aims of the study were:

- Investigate the role of the *lex2* loci (**Paper I and II**) acting as a glucosyl or galactosyltransferase.
- Establish the full LPS structure of *NTHi* 2019 and elucidate the gene responsible for adding *PEtn* at Hep III (**Paper III**).
- Elucidate the LPS structure of the invasive *NTHi* strain R2866 (**Paper IV**).



## RESULTS

---

### **PAPER I - A dual role for the *lex2* locus: identification of galactosyltransferase activity in non-typeable *Haemophilus influenzae* strain 1124 and 2019**

The aim of this study was to elucidate the function of the *lex2* locus (comprising of *lex2A* and *lex2B*), in *NTHi* strains which express a galactose residue in O-4 position on the conserved inner-core glucose (Glc I). Prior research has shown that the phase variable *lex2B* gene encode a glucosyltransferase responsible for the addition of glucose in O-4 position on Glc I in RM7004 and RM153 (Eagan) (Griffin R, 2003). Moreover, it has been shown that both *lex2A* and *lex2B* are needed for the transferase activity. The structural work performed in our lab and other labs have previously identified *NTHi* strains which express a galactose residue in the same position (Phillips NJ, 1992, Yildirim HH, 2005a). It therefore seemed plausible that the *lex2B* gene can encode either glucosyltransferases or galactosyltransferases. Indeed, these types of functional polymorphism features have been identified before, the *lpsA* gene, which is responsible for initiating OS elongation from Hep III can attach either glucose or galactose in  $\beta$ -(1 $\rightarrow$ 2)-linkage or  $\beta$ -(1 $\rightarrow$ 3)-linkage (Deadman ME, 2006a).

In order to test this hypothesis, the *lex2B* gene was inactivated in *NTHi* strains 1124 and 2019, producing mutant strains 1124*lex2* and 2019*lex2*. The LPS was extracted by the PCP-extraction method as described in the materials and methods section. The LPS was subjected to hydrozinolysis or delipidation to obtain OS or LPS-OH material. ESI-MS data on both OS and LPS-OH material indicated a heterogeneous mixture of glycoforms, in particular for 2019*lex2* which showed glycoforms of higher compositions than expected (Hex<sub>2-5</sub>·HexNAc<sub>0-1</sub>·Hep<sub>3</sub>·PEtn<sub>2-3</sub>). 1124*lex2* indicated two major compositions in the OS sample (PCho·Hex<sub>4</sub>·Hep<sub>3</sub>·PEtn<sub>2</sub>·AnKdo-ol and PCho·HexNAc·Hex<sub>4</sub>·Hep<sub>3</sub>·PEtn<sub>2</sub>·AnKdo-ol) which also were present in the O-deacetylated LPS (LPS-OH) sample.

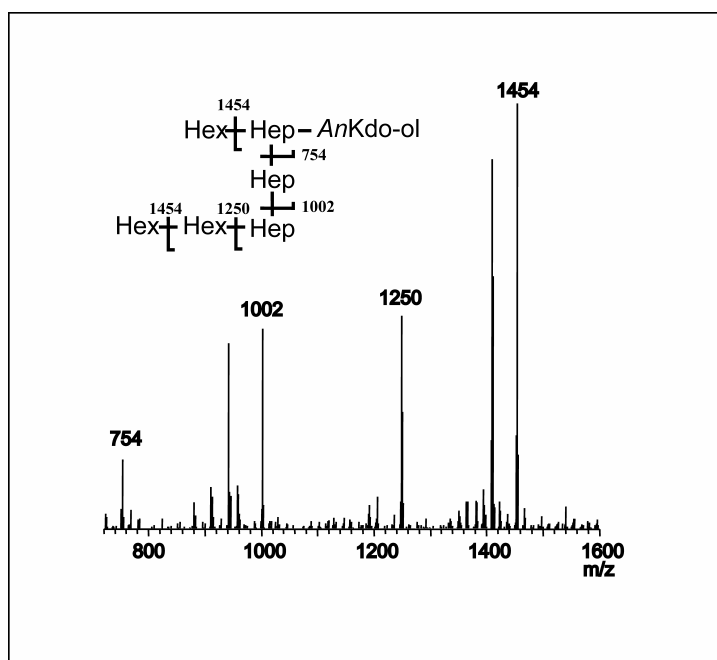
In order to examine the expression at Hep I, OS material of 1124*lex2* was dephosphorylated, permethylated and subjected to HPLC-ESI-MS<sup>n</sup> experiments. Briefly, the fullscan spectra displayed four glycoforms (Hex<sub>2</sub>, *m/z* 1468, Hex<sub>3</sub>, *m/z* 1672, Hex<sub>4</sub>, *m/z* 1876 and Hex<sub>4</sub>·HexNAc, *m/z* 2121) indicating terminal glucose at Hep I while Hep III could demonstrate a globotetraose epitope or truncated versions thereof.

To further strengthen these observations we conducted a full structural elucidation of OS material from 1124*lex2* by NMR. Interresidue NOE connectivities between Glc I H-1/Hep I H-4,6 established the sequence  $\beta$ -D-Glcp-(1 $\rightarrow$ 4)-L- $\alpha$ -D-Hep-(1 $\rightarrow$ ) thus further confirming our suggested role of *lex2B* in *NTHi* 1124.

Dephosphorylated and permethylated OS material from 2019*lex2* also indicated a mixtures of glycoforms in which the Glc I residue was found to be terminal.

## RESULTS

Furthermore, a glycoform at  $m/z$  2325 (Hex<sub>5</sub>·HexNAc) was determined to express a Hex-HexNAc-Hex-Hex epitope from Hep I. This is a result of *en bloc* addition controlled by the *hmg* locus as discussed in introduction, not related to the *lex2B* function. Past structural studies on NTHi 2019 (Gaucher SP, 2000, Phillips NJ, 1992) have focused primarily on expression from Hep I and Hep II, whereas the Hep III always has been indicated as a terminal residue. Our data indicates extension by two hexose residues from Hep III (**Figure 17**), this is further discussed and examined in **Paper III**.



**Figure 17.** HPLC-ESI-MS<sup>n</sup> experiments performed on dephosphorylated and permethylated OS material from 2019/*lex2*. The figure depicts the MS<sup>2</sup> spectra obtained from  $m/z$  1672 (Hex<sub>2</sub>·Hep<sub>3</sub>·AnKdo-ol). Included are the fragmentations for one of the glycoforms with a disaccharide branch at Hep III.

## RESULTS

---

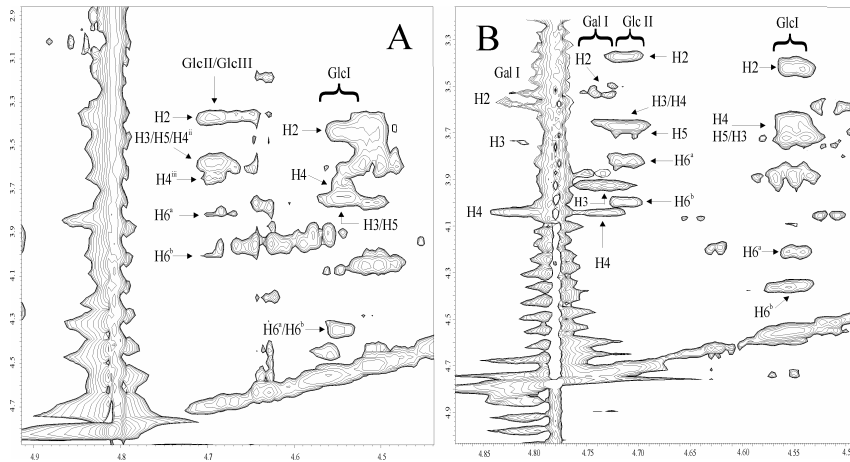
### **PAPER II - *Lex2B*, a phase-variable glycosyltransferase, adds either a glucose or a galactose to *Haemophilus influenzae* lipopolysaccharide.**

This study is a continuation of the investigation of the functions and genomic background of the *lex2* locus. This was studied by using a selection of *NTHi* strains obtained from patients with otitis media (Cody 2003). The *lex2* loci was first identified as a phase-variable LPS biosynthetic locus in a serotype b strain, strain DL42 (Jarosik GP, 1994) and controls oligosaccharide extension from the conserved glucose residue in the inner-core. This extension can continue by either a glucose residue or galactose (**Paper I**) depending on the amino acid sequence of the gene. This paper focuses upon the structural characterizations of two transformant strains of *NTHi* 1008 and 1124. The transformant strains were designated as follows: *NTHi* 1008 was transformed with the *lex2* locus from *NTHi* 1124 (a strain that normally adds galactose), the resulting strain was then designated 1008*lex2Gal*<sub>1124</sub>. The other transformant strain is therefore called 1124*lex2Glc*<sub>1008</sub> with the *lex2* locus from 1008 that normally express glucose.

HPLC-ESI-MS<sup>n</sup> experiments on dephosphorylated and permethylated OS material of 1008*lex2Gal*<sub>1124</sub> indicated a mixture of glycoforms ranging from Hex<sub>1</sub> to Hex<sub>6</sub>. A clear difference was apparent as compared to wild-type 1008 which expressed a Hex<sub>7</sub> glycoform as a major component of the LPS (Schweda, data not published). A successful replacement of Glc→Gal in *NTHi* 1008 would exclude a hexose residue from the epitope presented at Hep I. Indeed, trisaccharide epitopes were evidenced at Hep I and Hep III in the Hex<sub>6</sub>·Hep<sub>3</sub>·AnKdo-ol glycoform as shown by HPLC-ESI-MS<sup>n</sup>. In the case of Gal→Glc replacement, data obtained from HPLC-ESI-MS<sup>n</sup> experiments from OS1124*lex2Glc*<sub>1008</sub> indicated glycoforms corresponding to Hex<sub>3</sub>, Hex<sub>4</sub>, Hex<sub>5</sub>, Hex<sub>4</sub>·HexNAc, Hex<sub>6</sub>, and Hex<sub>7</sub>. The Hex<sub>7</sub>·Hep<sub>3</sub>·AnKdo-ol glycoform was shown to express a tetrasaccharide epitope at Hep I and a trisaccharide at Hep III. The inclusion of an additional hexose residue in the tetrasaccharide epitope at Hep I indicated a successful transformation, as this was not present in the wild-type strain (Yildirim HH, 2005a). The two transformant strains were also studied by NMR in order to conclude on the nature of the hexose identities. **Figure 18** shows the TOCSY spectra obtained from both transformant strains; from these experiments it was possible, in combination with NOESY experiments, to conclude that both transformations were successful.

Furthermore, this study includes the sequencing of the *lex2* loci from several *NTHi* strains which enabled us to compare the expected hexose expressed compared to structural studies performed. Through this comparison, it was found that the specificity to add either glucose or galactose depends on one single amino acid at position 157 in the *Lex2B* sequence. 8 of the 25 strains examined contained a threonine (T) at position

## RESULTS



**Figure 18.** TOCSY experiment (180 ms spinlock) performed on deacylated OS material from (A): 1124*lex2*Glc<sub>1008</sub> and (B): 1008*lex2*Gal<sub>1124</sub>. Glc II is situated at O-2 position on Hep III, Glc III and Gal I are linked to Glc I in O-4 position.

157 correlating with galactosyltransferase activity, while the remaining strains contained alanine (A), consistent with glucosyltransferase activity. In each of the strains, the hexose expressed is in agreement with the structural data. Six of these strains were selected for further genetic studies, three which were expected to express galactose, and three which express glucose residues linked to Glc I.

The six selected strains were mutated in two distinct ways, either the *lex2* gene was mutated (see **Paper I**) or alleles were switched between strains. The mutated strains were tested with a LPS-specific monoclonal antibody which binds digalactoside epitopes. These epitopes can be expressed at Hep I and dependant upon on the *lex2* loci, it was therefore possible to roughly compare LPS expression patterns between wild-type strains and mutant strains through sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and MAb binding. Resistance to killing by normal human serum was tested in order to compare wild-type strains with strains containing a dysfunctional *lex2* gene. These bactericidal tests indicated that killing by serum was enhanced in strains containing a dysfunctional *lex2* gene; possibly by disabling them to express a digalactoside epitope at Hep I. The mechanism by which the *lex2A* reading frame affects the phase-variable expression of the *lex2B* phenotype is still not clear. This study has identified *lex2B* as a glycosyltransferase that contributes to LPS heterogeneity through both its phase-variable characteristics and allelic polymorphism.

## RESULTS

---

### PAPER III – Detailed structural features of lipopolysaccharide glycoforms in *NTHi* strain 2019

This study focuses on the *NTHi* 2019, a strain which has been pivotal for several studies in laboratories worldwide. *NTHi* 2019 was the first strain for which the major LPS structures were characterized in 1992 (Phillips NJ, 1992).

Since then, this strain has been used to study biofilm formation (Greiner LL, 2004, Hong W, 2007, Jurcisek J, 2005, Jurcisek JA, 2007) and adherence to the PAF receptor (Swords WE, 2000). The structural knowledge concerning this model strain is limited, to date only the major expressed lactose epitope at Hep I has been properly established, even though mass spectrometry experiments (Gaucher SP, 2000) have shown that *NTHi* 2019 can produce glycoforms with up to five hexoses and one hexoseamine residue. All these studies state that Hep III is terminal.

Strain 2019 belongs to a small subset of *NTHi* strains which express two *PEtn* residues (Månsson M, 2003a, Yildirim HH, 2005a) decorating its conserved inner-core. It was recently shown that the *lpt6* gene is responsible for incorporating one of these residues in O-6 position on Hep II (Wright JC, 2004) while the remaining gene is unknown.

Moreover, the study performed by Swords and coworkers also included the use of a *pgmB* mutant which abolishes phosphoglucosyltransferase activity (interconverts glucose-1-phosphate and glucose-6-phosphate). A mutant strain should therefore have very restricted available glucose and the LPS should be severely truncated with only inner-core heptoses presented in the LPS structure. Swords and coworkers noted that this mutant strain produced LPS with four *PEtn* residues, no positions were established (Swords WE, 2000).

A candidate *PEtn* transferase gene, *lpt3* was identified from a study of simple sequence repeats in the recently available *H. influenzae* complete genome sequence (Power PM, 2009) and showed high sequence homology to genes encoding *PEtn* transferases in other bacteria such as *N. meningitidis*.

This paper contains the structural characterizations of several mutant strains of *NTHi* 2019 (*2019pgmB*, *2019lex2* and *2019lpt3*) as well as wild-type analysis. Each of these analyses will be discussed briefly in their separate sections as follows below.

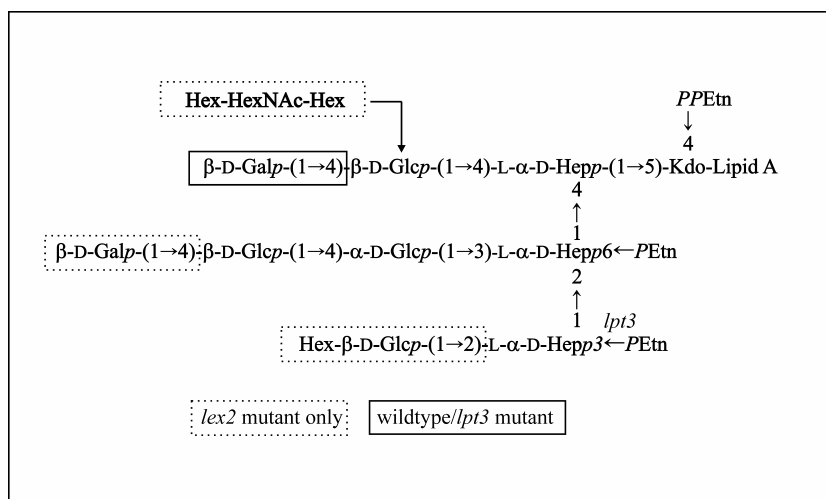
Structural determination of LPS expressed by **2019pgmB** evidenced the major component as the three heptoses as found in the conserved inner-core. However, ESI-MS experiments performed on both LPS-OH and OS material indicated a minor population with up to two hexoses. Compositional analysis and HPLC-ESI-MS<sup>n</sup> experiments identified trace amounts of a disaccharide unit linked to Hep I. It would be plausible that this epitope is a lactose unit as previously identified for *NTHi* 2019. Furthermore, the data obtained from these experiments indicated three *PEtn* residues, none of which were attached to Kdo or lipid A. CE-ESI-MS/MS data evidenced two of these residues linked to either Hep II or Hep III.

ESI-MS experiments performed on LPS-OH and OS material from **2019** and mutant strain **2019lpt3** indicated glycoforms ranging from Hex<sub>1</sub> to Hex<sub>5</sub>. The wild-type strain

## RESULTS

expressed minor amounts of compositions with one additional *PEtn* residue, these were absent in the *lpt3* mutant. HPLC-ESI-MS<sup>n</sup> experiments evidenced glycoforms summarized as Hex<sub>1,4</sub>:Hep<sub>3</sub>:AnKdo-ol. Two major isomeric structures were established, the first (Hex<sub>2</sub>:Hep<sub>3</sub>:AnKdo-ol) with a disaccharide unit at Hep I and the second (Hex<sub>3</sub>:Hep<sub>3</sub>:AnKdo-ol) with a monosaccharide at Hep I and a disaccharide branch at Hep II.

Major structures were elucidated by detailed <sup>1</sup>H, <sup>13</sup>C and <sup>31</sup>P NMR analyses on deacylated OS material using gradient chemical shift correlation techniques (COSY, TOCSY and HMQC). The sequence of the included saccharide units were established by intra- and inter-NOE connectivities. The absence of the additional *PEtn* units was apparent from HMQC experiments performed on 2019*lpt3* as compared to both wild-type strain and 2019*lex2* (discussed below). **Figure 19** indicate which elements were established in wild-type 2019, 2019*lpt3* and 2019*lex2*.



**Figure 19.** Structural elements identified from wild-type 2019 as well as 2019*lpt3* and 2019*lex2*. Solid boxes represent structures found in wild-type and *lpt3* mutant strain while dashed boxes represent unique elements found in 2019*lex2*.

Data obtained for analyses of LPS expressed by 2019*lex2* were consistent with previously published HPLC-ESI-MS<sup>n</sup> data (**Paper I**) in which the major components were identified as two Hex<sub>4</sub> glycoforms. The first isomer expressed monosaccharides from Hep I and Hep III while a disaccharide was evidenced substituting Hep II. The second isomer displayed a trisaccharide branch at Hep II while a monosaccharide was found at Hep I.

The additional *PEtn* residue was also present in the *lex2* mutant and was found to substitute Hep III in O-3 position as evidenced by NMR studies on deacylated OS

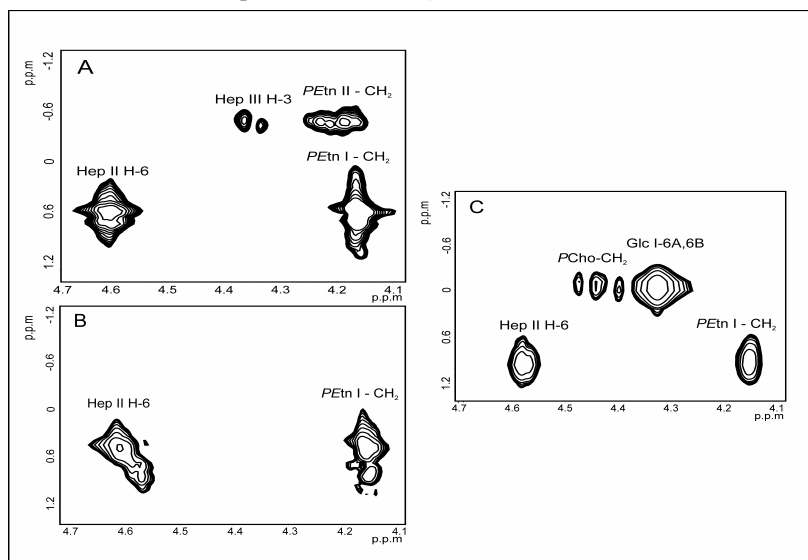
## RESULTS

material. This position has previously been identified as to carry *PEtn* residues in *NTHi* strains 1124 (Yildirim HH, 2005a) and 981 (Månsson M, 2003a). This study highlights an important possibility, that genes can affect each other. When the *Lex2* transferase was knocked out in 2019*lex2* it affected the expression controlled by *lpsA* allowing for elongations which previously never have been reported for this strain. This is an important observation that our current knowledge cannot explain and something which needs to be studied further.

Also the same mutant strain demonstrates gene competition, as O-4 substitution from Glc I is controlled by either *lex2* or the *hmg* locus. When *lex2* activity was removed it enabled us to detect the *en bloc* expressions that were not detected previously for this particular *NTHi* strain.

The phophoglucomutase mutant strain, 2019*pgmB* also gave clear indications for alternative ways in which the bacterium can express hexoses in its outer core, something which needs to be studied further. Since this particular mutant is used frequently in biofilm investigations, it is important to note that the expected loss of hexoses in the outer core might not be true. Our studies show that this mutant expresses trace amounts of a possible lactose epitope, an acceptor for Neu5Ac with vital roles in biofilm formation.

This investigation also identified the *lpt3* gene as being responsible for incorporation of *PEtn* in O-3 position at Hep III. Further studies were carried out on *NTHi* 981, a strain with 100% expression of this O-3 linked *PEtn*. The *lpt3* mutant of this strain indicated the same result (Hood, unpublished data) (**Figure 20**).



**Figure 20.**  $^{31}\text{P}$ - $^1\text{H}$  HMQC spectrum of deacylated OS material from (A) 2019*lex2* (B) 2019*lpt3* and (C) 981*lpt3*. The proposed cross peaks between *PEtn* and H-3 of Hep III, *PEtn* and H-6 of Hep II and *PCho* and H-6<sub>A,B</sub> of Glc I are indicated.

## RESULTS

---

### PAPER IV - Structural analysis of the lipopolysaccharide from invasive non-typeable *Haemophilus influenzae* strain R2866

NTHi strain R2866 is a clinical isolate obtained from a patient with invasive disease (Nizet V, 1996). Invasive disease caused by non-typeable strains of *Haemophilus influenzae* is uncommon, but some cases have been reported (Campos J, 2004, Cerquetti M, 2000, Cuthill SL, 1999, O'Neill JM, 2003). Because of this rare characteristic, we performed a full structural characterization of the LPS expressed by this strain to see if any features in the LPS could correlate to this behavior. The genome of strain R2866 has been sequenced recently (Erwin AL, 2005) which provides an outstanding opportunity to study novel LPS structures and the relationship to biosynthetic genes. Furthermore, R2866 belongs to a set of NTHi strains which display heptoses as part of the outer core (Lundström SL, 2008a, Månsson M, 2003a, Månsson M, 2003b). Recently, two heptosyltransferases were identified through homology to *Haemophilus ducreyi* (Post DM, 2007); the respective genes are now referenced to as *losB1* and *losB2*. Prior studies conducted by our lab have identified *losB1* as being the gene responsible for incorporation of D,D-Hep while *losB2* is responsible for L,D-Hep incorporation in the outer core, both in O-6 position of Glc I (Lundström SL, 2008a). R2866 contain the *losB2* loci and should therefore express L,D-Hep in its outer core. Furthermore, a presumed hexosyltransferase gene named *losA2* (Tullius MV, 2002) has been identified.

Two mutant strains were constructed in order to simplify the LPS expressed by this highly heterogeneous strain, the *lpsA* mutant effectively disables elongation possibilities from Hep III while the heptosyltransferase mutant enabled us to characterize epitopes related to O-4 substitution by the *lex2* loci from Glc I.

**Figure 21** shows all structural features as identified from wild-type strain as well as unique elements found in the two genetically defined mutant strains R2866*losB2* and R2866*lpsA*. The variable expression of carbohydrate epitopes presented by R2866 is staggering; overall 35 different isomeric forms were evidenced by HPLC-ESI-MS<sup>n</sup> experiments performed on the wild-type strains as well as the *lpsA* and *losB2* mutant strains.

HPLC-ESI-MS<sup>n</sup> experiments performed on OS material from wild-type R2866 indicated 15 different isomeric forms with compositions summarized as

Hex<sub>1-7</sub>·Hep<sub>3-4</sub>·HexNAc<sub>0-1</sub>·AnKdo-ol. The major glycoform expressed was shown to carry a trisaccharide epitope at Hep III while Glc I was substituted by a L- $\alpha$ -D-Hep $\beta$ -(1 $\rightarrow$  unit in O-6 position. Of particular interest was the substitution with PCho which was evidenced in all glycoforms expressed. The combination of CE-ESI-MS/MS and NMR experiments evidenced this residue to substitute L,D-Hep IV in O-7 position, a position which is novel. Methylation analysis also indicated minor amounts of a 4,6-

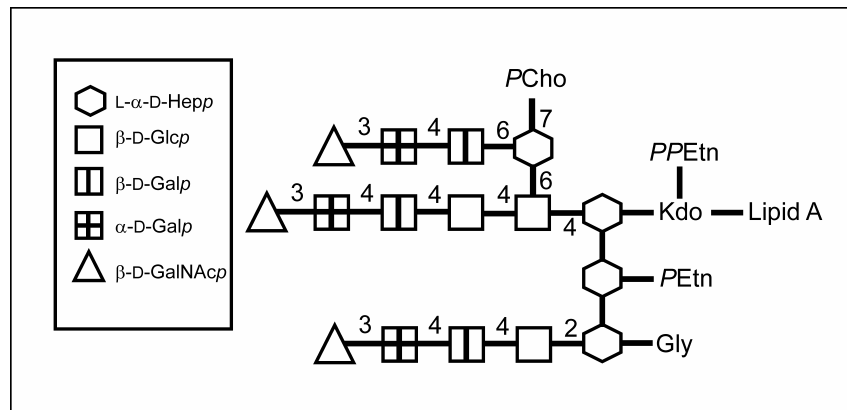


## RESULTS

substituted Glc residue (Glc I), typically seen for branched structures (Månsson M, 2003a). Characterization of the LPS expressed by R2866*losB2* evidenced a  $\beta$ -D-GalNAcp-(1→3)- $\alpha$ -D-Galp-(1→4)- $\beta$ -D-Galp-(1→4)- $\beta$ -D-Glcp-(1→ chain with attachment to O-4 position on Glc I with no indications of outer core heptoses as suggested by the *losB2* function. The expression of *PCho* was lost once the outer core heptose was lost. Moreover, a globoside epitope was evidenced to substitute Hep III as demonstrated in wild-type strain.

The *lpsA* mutant of R2866 indicated substitution on the outer core heptose in O-6 position, a feature which was not seen in the wild-type strain. This has previously been reported in other *NTHi* strains such as strain 1233 (Månsson M, 2003b) were it carried a  $\beta$ -D-GalNAcp-(1→3)- $\alpha$ -D-Galp-(1→4)- $\beta$ -D-Galp-(1→6 extension. The combination of methylation analysis, permethylation analysis and ESI-MS experiments performed on OS material from R2866*lpsA* provide substantial evidence that this epitope also is present in R2866.

This epitope could relate to the novel glycosyltransferase (*losA2*). As this gene also is present in *NTHi* 1233, it would thus be reasonable to assume that this gene adds  $\beta$ -D-Galp in O-6 position on outer core L,D-Hep. This theory will be examined by us in the near future.



**Figure 21.** Structural elements as identified from wild-type R2866 as well as genetically defined mutant strains R2866*lpsA* and R2866*losB2*.

## CONCLUSIONS AND FUTURE PERSPECTIVES

---

To explore the relationship between expressed structure and their corresponding biosynthetic LPS genes, we elucidated the structures of LPS from both wild-type and genetically defined mutant strains of *NTHi*. This enabled us to correlate biosynthetic gene functions with structural motifs identified by analytical techniques and methods such as mass spectrometry and NMR.

**Paper I** and **Paper II** demonstrate how powerful this approach is. In these papers we draw conclusions centered on the dual nature of the *lex2* gene's function.

**Paper III** demonstrates an example of the interplay between two biosynthetic genes, where inactivating *lex2* up regulates the expression by another gene, namely *lpsA*. In this paper we were able to prove that Hep III of *NTHi* 2019 also can carry further carbohydrate substitution, a feature which previously had never has been reported for this strain. The phosphoethanolamine transferase gene *lpt3* was also characterized as being responsible for adding *PEtn* in O-3 position on Hep III.

**Paper IV** is one of the first structural characterizations of LPS from a *NTHi* strain which is able to cause invasive disease. A novel O-7 substitution of *PCho* on the outer core L,D-Hep IV was identified. The characterization of LPS expressed by R2866*lpsA* indicated glycoforms in which Hep IV was further substituted in O-6 position by Gal. This could relate to the newly discovered glycosyltransferase genes *losA*.

As demonstrated for the *lpsA* gene, the *lex2* loci described in **Paper I/II** possesses a dual-nature which enables the same gene to produce either a glucose extension or galactose extension from Glc I. As to date, the following epitopes which are initiated by *lex2* have been identified:  $\beta$ -D-GalNAcp-(1 $\rightarrow$ 3)- $\alpha$ -D-Galp-(1 $\rightarrow$ 4)- $\beta$ -D-Galp-(1 $\rightarrow$ 4),  $\beta$ -D-GalNAcp-(1 $\rightarrow$ 3)- $\alpha$ -D-Galp-(1 $\rightarrow$ 4)- $\beta$ -D-Galp-(1 $\rightarrow$ 4)- $\beta$ -D-Glc-(1 $\rightarrow$ 4) and truncated versions thereof. Two of these truncated versions include the sialic acid acceptor lactose and the digalactoside epitope. Sialylated epitopes and expression of digalactose structures are known to contribute to increased virulence, thus making these findings biologically relevant.

A key feature of the *lex2* gene that differs from *lpsA* is its phase-variable nature which makes it possible to switch off/on its expression depending on the requirements presented by the host environment.

A single key amino acid in position 157 determines whether a galactosyl or glucosyltransferase activity will be present, alanine (A) correlates with glucose while threonine (T) relates to galactose. The biological effect of galactose expressions compared to glucose expression is unclear. Both allow for further extensions by digalactoside epitopes and present lactose acceptors for sialic acid. It would therefore be interesting to study if the inclusion or exclusion of a hexose residue in the epitope relate to any biologically relevant behavior.

Prior studies have indicated that the glycosyltransferase encoded by *lex2B* is functional only of *lex2A* is expressed; this occurs when there are 3+3n repeats in the *lex2A*

## CONCLUSIONS AND FUTURE PERSPECTIVES

---

sequence. In contrast to these observation is the fact that *lex2A* is considered “off” in NTHi 2019 (**Paper II**), due structural data clearly show that expression is present.

More

experiments are needed in order to fully understand how *lex2A* works and what is required for full *lex2B* functionality. This is another example of allelic variation that has been reported for *H. influenzae*, most likely the phase variable gene *lic2B* will display similar characteristics as *lex2B* once properly analyzed.

In **Paper III** we demonstrate that Hep III can be further substituted by a  $\beta$ -D-Glcp-(1→2) residue which allows for further elongations by  $\beta$ -D-GalNAcp-(1→3)- $\alpha$ -D-Galp-(1→4)- $\beta$ -D-Galp-(1→4) or  $\alpha$ -Neu5Ac-(2→8)- $\alpha$ -Neu5Ac-(2→3)- $\beta$ -D-Galp-(1→4), something which previously never has been seen for NTHi 2019. The epitope expressed from Hep II was also fully characterized.

We also show that the *pmgB* mutant still is able to express a disaccharide epitope at Hep I, a epitope which might function as an acceptor for Neu5Ac. All of these structural data are important when one debates the topics such as biofilm and to be able to make correct assumptions on its formation, we believe that the LPS structure of the model strain should be completely understood.

Moreover, the phosphoethanolamine transferase gene *lpt3* has been identified and its function has been established. In the future it would be interesting to study the implication of these residues in correlation to pathogenicity.

In **Paper IV** we characterized a NTHi strain which causes invasive disease in otherwise healthy patients.

R2866 is one of the first invasive NTHi strains for which the LPS structure has been elucidated and some features might be relevant to its special pathogenic behavior: the PCho found in O-7 position on Hep IV, multiple digalactoside epitopes and possibility to express a multitude of Neu5Ac related epitopes.

Past studies have indicated that the position of PCho in the LPS influences how well CRP can initiate clearance by the immune response. O-7 substitution of outer core L,D-Hep is a novel position, most likely not recognized as well by CRP as classical O-6 substitution on Glc I. A detailed investigation of the *lic1* sequence is needed in order to find variants responsible for incorporation PCho in this new novel position.

Expression of two digalactoside epitopes at once provides special benefit in order to survive killing by normal human serum, this is also the case for sialylated epitopes. These factors can contribute to the bacterium’s ability to persist, colonize and survive within the host.

The recently published genome sequences for R2846 and R2866 provided an outstanding opportunity to study previously uncharacterized biosynthetic genes related to LPS expression. Two new heptosyltransferases, *losB1* and *losB2*, were recently identified as being responsible for the incorporation of D,D-Hep and L,D-Hep to the outer core of NTHi.

## CONCLUSIONS AND FUTURE PERSPECTIVES

---

In conjunction with these heptosyl related genes came two glycosyltransferase genes, *losA1* and *losA2*. When the *losB/losA* genes are present as a pair, this excludes the *lic2C* and *lic2B* genes and thus prevent any elongations from Hep II. We suggest that *losA1/losA2* are the prime candidates to add either glucose or galactose onto the outer core heptoses produced by *losB1* and *losB2*. The genome sequence of R2866 evidenced the *losB2/losA2* genes to be present in the strain.

Interestingly, the elucidation of the *lpsA* mutant of R2866 showed signs of galactose substitution in O-6 position on the outer core heptose, something which previously has been reported for *NTHi* strain 1233. *NTHi* R2866*lpsA* and 1233*lpsA* have many structural features in common, moreover they both contain the *losB2/losA2* gene pair. It is therefore plausible that *losA2* controls the addition of  $\beta$ -D-Galp to O-6 position on L,D-Hep.

Future studies are needed in order to fully understand the nature of the *losB/losA* genes, a matter which is complicated by the fact that inactivating any of the *losA* genes leads to the automatic inactivation of the related *losB* gene. This was solved in a highly sophisticated way when Deadman and coworkers (Deadman, unpublished data) constructed the R2846*losA1/losB1*+ mutant. The *losA1* gene was inactivated after which the *losB1* gene was reintroduced back into the strain. Indeed, these data indicates that *losA1* controls the addition of  $\beta$ -D-Glcp in O-4 position on external D,D-Hep.

## POPULÄRVETENSKAPLIG SAMMANFATTNING

---

En av de vanligaste orsakerna till ett barns första besök hos sin läkare är någon form av öroninflammation. Termen ”öronbarn” är något som de flesta av oss kan relatera till, antingen eftersom man själv har barn med återkommande problem eller att man själv hade/har problem med öronen. I denna avhandling undersöker vi en av bakterierna som leder till dessa besvär, bakterien *Haemophilus influenzae*.

Öroninflammation må verka oskyldigt vid första anblicken, men bakterien kan även orsaka betydligt värre åkommor om den lyckas ta sig ut i blodet. Om bakterien når hjärnan kan den orsaka hjärnhinneinflammation som kan resultera i dödsfall.

I denna avhandling har vi studerat molekyler som består av socker som sitter på bakterieytan av *Haemophilus influenzae*, dessa kallas lipopolysackarider. Dom här molekylerna påverkar bakteriens förmåga att försvara sig mot människokroppen (immunförsvaret) och dess möjlighet att orsaka sjukdom. Detta beror på att dom sitter längst ut på bakteriens utsida och exponeras mot miljön inne i våra kroppar, de interagerar.

Genom arbetet som genomförs i vår grupp har vi kunnat bestämma hur dessa molekyler ser ut i olika så kallade bakteriestammar av *Haemophilus influenzae*. Den här informationen kan användas för att i framtiden framställa vaccin. Tidigare försök med detta på en typ av *Haemophilus influenzae* har varit mycket lyckade, tyvärr fungerar detta vaccin inte på den variant av bakterien som denna avhandling handlar om.

Alla vi människor består av miljarder celler, och på många av dessa cellers yta finns sockermolekyler, till exempel på våra röda blodkroppar. Sockermolekylerna på *Haemophilus influenzae*'s yta är smart placerade, eftersom de härmar de sockermolekyler som finns på våra andra celler. Detta gör att bakterien kan gömma sig inne i kroppen som en kameleont, kroppens försvar ser inte något som är annorlunda.

Som allting annat i vår fantastiska kropp styrs även bakteriens olika delar av gener. Gener leder till att saker produceras eller slutar produceras, till exempel dessa lipopolysackarider. I denna avhandling har flera gener som har att göra med sockermolekylerna på *Haemophilus influenzae*'s yta studerats. Ökad kunskap om en specifik gen kan ge information om hur bakterien orsakar sjukdom.

En av dessa gener har studerats i detalj i denna avhandling, den kallas för *lex2*. Denna gen påverkar i stor grad hur pass bra bakterien är på att överleva i vår kropp. Om en bakterie är bra på att överleva länge är det naturligt att den förr eller senare orsakar besvär.

Förutom detta har vi också studerat en speciell bakteriestam av *Haemophilus influenzae* som kan orsaka de farligare sjukdomarna som hjärnhinneinflammation oftare. Genom att bestämma hur sockermolekylerna ser ut hos denna stam hoppas vi på att kunna dra slutsatser om varför just denna stam är farligare än andra.

## POPULÄRVETENSKAPLIG SAMMANFATTNING

---

Allt detta har vi gjort genom att använda olika instrument och metoder som gör det möjligt att titta på strukturen av en stor molekyl ner till minsta atom. En av de viktigaste metoderna att studera strukturer kallas masspektrometri. Masspektrometri fungerar genom att man tar hela sin molekyl och slår sönder den i små eller stora bitar med hjälp av energi. Varje molekyl går sönder på sitt speciella sätt, man kan säga att dom har ett eget fingeravtryck. Genom att studera dessa fingeravtryck kan man i många fall säga hur molekylerna såg ut innan den gick sönder.

Olika gener och deras funktion har också studerats, bland annat genen *lex2*. Detta kan man göra genom att inaktivera genen ifråga och sedan se om något i molekylens struktur ändras. För att göra detta krävs att man redan vet något om hur sockermolekylerna ser ut i deras vanliga tillstånd, alltså när alla gener fungerar som de ska. I denna avhandling har vi tittat på bakteriestammar där alla gener fungerar som de ska, men även stammar där vissa gener har inaktiverats.

Vi hoppas att vår forskning i framtiden kan användas för vaccinutveckling och på så sätt underlätta för alla dessa "öronbarn".

## ACKNOWLEDGEMENTS

---

To my main supervisor **Elke Schweda** for skillful supervision during my time in your group. In particular, I am thankful for the amount of time you spend with your Ph D students and for always keeping your door open. I wish to thank you for allowing me to perform both my exam work and Ph D studies in your group and for introducing me to a challenging and interesting field of research. You have always listen to my troubles, both work-related and personal issues and given me great advice.

To **Derek Hood** my co-supervisor for helping me understand the wider picture in our research. We have had great discussion which I have truly enjoyed which have further sparked my interest in microbiology. To **Mary Deadman** for great collaborations with the papers included in this thesis. You are a very kind soul, something which is needed in the research world. To **Richard Moxon** for his incredible knowledge in the field and for providing an exciting and eventful meeting in Oxford.

To **Jim Richards** for great scientific discussions and for many great christmas tables. Its safe to say that everybody in our group have enjoyed every visit you have made to Sweden. I'm very thankful for all your ideas that have lead to great publications and for all the input that have made our publications even better.

To **Jianjun Li** for contributing with CE-ESI-MS/MS data and for great collaborations in the papers included in this thesis.

To **Alice Erwin** and **Arnold Smith** for sharing the genome sequence of R2866 and for providing the strain.

To the **Finnish Otitis Media Study Group** for providing numerous *NTHi* strains used in this thesis.

To **Håkan Yildirim** for helping me during my exam work. You made me think in a scientific way and made us all laugh out load many times. By the way, I guess you won the bet, I did marry within two years!

To **Ivan Mikhail** for great times, the pubs at Södertörn's were particulary interesting.

To **Susanna Lundström**. You were easy to talk to, and always gave advice that was well thought through. Being able to both be a serious researcher and a relaxed friend is sometimes difficult, but you did prove that its possible. I will miss our walks after lunch!

To **Varvara Vitiazeva** for her extremely talent for detail and skill with NMR. Many times you helped me solve things that I got stuck on by just glancing at the problem, very impressive! I am also grateful for all advice dealing with life itself, there is a lot of good in you.

## ACKNOWLEDGEMENTS

---

To **Brigitte Twelkmeyer** for lightening up my days in the lab. I loved our times at the pub, pulling down beer talking about nothing and everything at once. Your positive view-point of life is good for the people around you! Also you brought us closer to other groups by planning the christmas table.

To **Anna-next-Door**, our office neighbour. Thank you for being kind and making me relax a bit during intensive workdays.

To all the people at **Clinical Research Center** and **Laboratory Medicine**, for a lot of fun times during meeting and conferances.

To all the people of **Experimental Cancer Medicine**. We had so much fun during christmas tables and dissertations. I've learnt quite a lot about Oman, Egypt and Iran and was presented to food and drinks that I never had before, delicious!

To **the administration**, without you guys work would have been much more difficult. In particular I wish to thank **Hanna Eriksson** for helping me organize everything around my defense and for providing many friendly advices.

To Södertörns University Collage for providing an excellent education and for partially financing my Ph D studies.

To **my parents** for teaching me the importance of setting high standards for oneself. Without your kind nurture and guidance I would never have been where I am today. I wish to thank **my father** for his constant concern and extremely generous attitude during my school years. I wish to thank **my mother** for understanding me, putting faith in me and always putting my needs first. Without you, this would not have been possible!

To my **parents-in-law, Athanassios and Fotini** for letting me into their big fat greek family. After I got wed to your wonderful daughter, the amounts of relatives I had increased by 1000%. I wish to thank you for many lovely summers in Greece and for all eventful Sunday dinners (and all other dinners too of course).

To **my wife, Pinelopi**. If I could marry you twice I would without the blink of an eye. You have brought love into my life in a way I never thought was possible. Without your constant support, guidance and care I would be less of a man and with half of a heart. Words cannot discribe what I feel for you, Σ'αγαπώ αγάπη μου!

To **my son** for being an absolute wonder. When people discribe how it feels to be a father I never did understand truely until now. You brought joy into my life in ways which are impossible to describe.



## REFERENCES

---

- Alexander, C. and Rietschel, E. T. (2001) Bacterial lipopolysaccharides and innate immunity, *J Endotoxin Res.* 7, 167-202.
- Armbruster, C. E., Hong, W., Pang, B., Dew, K. E., Juneau, R. A., Byrd, M. S., Love, C. F., Kock, N. D. and Swords, W. E. (2009) LuxS promotes biofilm maturation and persistence of nontypeable haemophilus influenzae in vivo via modulation of lipooligosaccharides on the bacterial surface, *Infect Immun.* 77, 4081-4091.
- Auzanneau, F.-I., Charon, D. and Szabó, L. (1991) Phosphorylated Sugars. Part 27. Synthesis and Reactions, in Acidic Medium, of 5-O-Substituted Methyl 3-Deoxy- $\alpha$ -D-manno-oct-2-ulopyranosidonic Acid 4-Phosphates, *J. Chem. Soc. Perkin Trans.*, 1, 509-517.
- Barenkamp, S. J. and St Geme, J. W., 3rd. (1994) Genes encoding high-molecular-weight adhesion proteins of nontypeable Haemophilus influenzae are part of gene clusters, *Infect Immun.* 62, 3320-3328.
- Barenkamp, S. J. and St Geme, J. W., 3rd. (1996) Identification of a second family of high-molecular-weight adhesion proteins expressed by non-typable Haemophilus influenzae, *Mol Microbiol.* 19, 1215-1223.
- Barry, B., Gehanno, P., Blumen, M. and Boucot, I. (1994) Clinical outcome of acute otitis media caused by pneumococci with decreased susceptibility to penicillin, *Scand J Infect Dis.* 26, 446-452.
- Bauer, S. H., Månsson, M., Hood, D. W., Richards, J. C., Moxon, E. R. and Schweda, E. K. (2001) A rapid and sensitive procedure for determination of 5-N-acetyl neuraminic acid in lipopolysaccharides of *Haemophilus influenzae*: a survey of 24 non-typeable *H. influenzae* strains, *Carbohydr Res.* 335, 251-260.
- Blakeney, A. B. and Stone, B. A. (1985) Methylation of carbohydrates with lithium methylsulphinyl carbanion, *Carbohydrate Research.* 140, 319-324.
- Bouchet, V., Hood, D. W., Li, J., Brisson, J. R., Randle, G. A., Martin, A., Li, Z., Goldstein, R., Schweda, E. K., Pelton, S. I., Richards, J. C. and Moxon, E. R. (2003) Host-derived sialic acid is incorporated into *Haemophilus influenzae* lipopolysaccharide and is a major virulence factor in experimental otitis media, *Proc Natl Acad Sci U S A.* 100, 8898-8903.
- Buttery, J. and Moxon, E. R. (2002) Capsulate bacteria and the lung, *Br Med Bull.* 61, 63-80.
- Campagnari, A. A., Gupta, M. R., Dudas, K. C., Murphy, T. F. and Apicella, M. A. (1987) Antigenic diversity of lipooligosaccharides of nontypable *Haemophilus influenzae*, *Infect Immun.* 55, 882-887.
- Campos, J., Hernando, M., Roman, F., Perez-Vazquez, M., Aracil, B., Oteo, J., Lazaro, E. and de Abajo, F. (2004) Analysis of invasive Haemophilus influenzae infections after extensive vaccination against H. influenzae type b, *J Clin Microbiol.* 42, 524-529.
- Cerquetti, M., Ciofi degli Atti, M. L., Renna, G., Tozzi, A. E., Garlaschi, M. L. and Mastrantonio, P. (2000) Characterization of non-type B Haemophilus influenzae strains isolated from patients with invasive disease. The HI Study Group, *J Clin Microbiol.* 38, 4649-4652.
- Cody, A. J., Field, D., Feil, E. J., Stringer, S., Deadman, M. E., Tsolaki, A. G., Gratz, B., Bouchet, V., Goldstein, R., Hood, D. W. and Moxon, E. R. (2003) High rates of recombination in otitis media isolates of non-typeable Haemophilus influenzae, *Infect Genet Evol.* 3, 57-66.
- Cox, A. D., Hood, D. W., Martin, A., Makepeace, K. M., Deadman, M. E., Li, J., Brisson, J. R., Moxon, E. R. and Richards, J. C. (2002) Identification and structural characterization of a sialylated lacto-N-neotetraose structure in the lipopolysaccharide of Haemophilus influenzae, *Eur J Biochem.* 269, 4009-4019.
- Cox, A. D., Masoud, H., Thibault, P., Brisson, J. R., van der Zwan, M., Perry, M. B. and Richards, J. C. (2001) Structural analysis of the lipopolysaccharide from the

- nontypable *Haemophilus influenzae* strain SB 33, *Eur J Biochem.* 268, 5278-5286.
- Cuthill, S. L., Farley, M. M. and Donowitz, L. G. (1999) Nontypable *Haemophilus influenzae* meningitis, *Pediatr Infect Dis J.* 18, 660-662.
- Danan, A., Mondage, M., Sarfati, S. R. and Szabó, P. (1982) Synthesis and Behaviour under Acidic Conditions of 2-Deoxy-D-arabino-hexopyranose and 3-Deoxy-2-ketoaldonic Acids Bearing *O*-Phosphono or *O*-Glycosyl Substituents at Position  $\beta$  to the Carbonyl Function, *J. Chem. Soc. Perkin Trans.*, I, 1275-1282.
- Deadman, M. E., Lundström, S. L., Schweda, E. K., Moxon, E. R. and Hood, D. W. (2006a) Specific amino acids of the glycosyltransferase LpsA direct the addition of glucose or galactose to the terminal inner core heptose of *Haemophilus influenzae* lipopolysaccharide via alternative linkages, *J Biol Chem.* 281, 29455-29467.
- Deadman, M. E., Lundström, S. L., Schweda, E. K., Moxon, E. R. and Hood, D. W. (2006b) Specific amino acids of the glycosyltransferase LpsA direct the addition of glucose or galactose to the terminal inner core heptose of *Haemophilus influenzae* lipopolysaccharide via alternative linkages, *J Biol Chem.* 281, 29455-29467.
- Domon, B. and Costello, C. E. (1988) Structure elucidation of glycosphingolipids and gangliosides using high-performance tandem mass spectrometry, *Biochemistry.* 27, 1534-1543.
- Duus, J., Gottfredsen, C. H. and Bock, K. (2000) Carbohydrate structural determination by NMR spectroscopy: modern methods and limitations, *Chem Rev.* 100, 4589-4614.
- Erwin, A. L., Nelson, K. L., Mhlanga-Mutangadura, T., Bonthuis, P. J., Geelhood, J. L., Morlin, G., Unrath, W. C., Campos, J., Crook, D. W., Farley, M. M., Henderson, F. W., Jacobs, R. F., Muhlemann, K., Satola, S. W., van Alphen, L., Golomb, M. and Smith, A. L. (2005) Characterization of genetic and phenotypic diversity of invasive nontypeable *Haemophilus influenzae*, *Infect Immun.* 73, 5853-5863.
- Erwin, A. L. and Smith, A. L. (2007) Nontypeable *Haemophilus influenzae*: understanding virulence and commensal behavior, *Trends Microbiol.* 15, 355-362.
- Farley, M. M., Stephens, D. S., Mulks, M. H., Cooper, M. D., Bricker, J. V., Mirra, S. S. and Wright, A. (1986) Pathogenesis of IgA1 protease-producing and -nonproducing *Haemophilus influenzae* in human nasopharyngeal organ cultures, *J Infect Dis.* 154, 752-759.
- Figueira, M. A., Ram, S., Goldstein, R., Hood, D. W., Moxon, E. R. and Pelton, S. I. (2007) Role of complement in defense of the middle ear revealed by restoring the virulence of nontypeable *Haemophilus influenzae* siaB mutants, *Infect Immun.* 75, 325-333.
- Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J. F., Dougherty, B. A., Merrick, J. M. and et al. (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd, *Science.* 269, 496-512.
- Fox, K. L., Cox, A. D., Gilbert, M., Wakarchuk, W. W., Li, J., Makepeace, K., Richards, J. C., Moxon, E. R. and Hood, D. W. (2006) Identification of a bifunctional lipopolysaccharide sialyltransferase in *Haemophilus influenzae*: incorporation of disialic acid, *J Biol Chem.* 281, 40024-40032.
- Fox, K. L., Li, J., Schweda, E. K., Vitiaveva, V., Makepeace, K., Jennings, M. P., Moxon, E. R. and Hood, D. W. (2007) Duplicate Copies of Lic1 Direct the Addition of Multiple Phosphocholine Residues in the Lipopolysaccharide of *Haemophilus Influenzae*, *Infect Immun.*
- Fox, K. L., Yildirim, H. H., Deadman, M. E., Schweda, E. K., Moxon, E. R. and Hood, D. W. (2005) Novel lipopolysaccharide biosynthetic genes containing tetranucleotide repeats in *Haemophilus influenzae*, identification of a gene for adding O-acetyl groups, *Mol Microbiol.* 58, 207-216.
- Foxwell, A. R., Kyd, J. M. and Cripps, A. W. (1998) Nontypeable *Haemophilus influenzae*: pathogenesis and prevention, *Microbiol Mol Biol Rev.* 62, 294-308.

- Galanos, C., Lüderitz, O. and Westphal, O. (1969) A new method for the extraction of R lipopolysaccharides, *Eur J Biochem.* 9, 245-249.
- Gaucher, S. P., Cancilla, M. T., Phillips, N. J., Gibson, B. W. and Leary, J. A. (2000) Mass spectral characterization of lipooligosaccharides from *Haemophilus influenzae* 2019, *Biochemistry.* 39, 12406-12414.
- Gerwig, G. J., Kamerling, J. P. and Vliegthart, J. F. (1979) Determination of the absolute configuration of mono-saccharides in complex carbohydrates by capillary G.L.C, *Carbohydr Res.* 77, 1-7.
- Giebink, G. S. (1981) The pathogenesis of pneumococcal otitis media in chinchillas and the efficacy of vaccination in prophylaxis, *Rev Infect Dis.* 3, 342-353.
- Gilsdorf, J. R. (1998) Antigenic diversity and gene polymorphisms in *Haemophilus influenzae*, *Infect Immun.* 66, 5053-5059.
- Greiner, L. L., Watanabe, H., Phillips, N. J., Shao, J., Morgan, A., Zaleski, A., Gibson, B. W. and Apicella, M. A. (2004) Nontypeable *Haemophilus influenzae* strain 2019 produces a biofilm containing N-acetylneuraminic acid that may mimic sialylated O-linked glycans, *Infect Immun.* 72, 4249-4260.
- Griffin, R., Bayliss, C. D., Herbert, M. A., Cox, A. D., Makepeace, K., Richards, J. C., Hood, D. W. and Moxon, E. R. (2005) Digalactoside expression in the lipopolysaccharide of *Haemophilus influenzae* and its role in intravascular survival, *Infect Immun.* 73, 7022-7026.
- Griffin, R., Cox, A. D., Makepeace, K., Richards, J. C., Moxon, E. R. and Hood, D. W. (2003) The role of *lex2* in lipopolysaccharide biosynthesis in *Haemophilus influenzae* strains RM7004 and RM153, *Microbiology.* 149, 3165-3175.
- Hakomori, S. (1964) A Rapid Permethylation of Glycolipid, and Polysaccharide Catalyzed by Methylsulfinyl Carbanion in Dimethyl Sulfoxide, *J Biochem.* 55, 205-208.
- Harrison, A., Dyer, D. W., Gillaspay, A., Ray, W. C., Mungur, R., Carson, M. B., Zhong, H., Gipson, J., Gipson, M., Johnson, L. S., Lewis, L., Bakaletz, L. O. and Munson, R. S., Jr. (2005) Genomic sequence of an otitis media isolate of nontypeable *Haemophilus influenzae*: comparative study with *H. influenzae* serotype d, strain KW20, *J Bacteriol.* 187, 4627-4636.
- Helander, I. M., Lindner, B., Brade, H., Altmann, K., Lindberg, A. A., Rietschel, E. T. and Zähringer, U. (1988) Chemical structure of the lipopolysaccharide of *Haemophilus influenzae* strain I-69 Rd-/b+. Description of a novel deep-rough chemotype, *Eur J Biochem.* 177, 483-492.
- Hellerqvist, C. G. (1990) Linkage analysis using Lindberg method, *Methods Enzymol.* 193, 554-573.
- Herriott, R. M., Meyer, E. M. and Vogt, M. (1970) Defined nongrowth media for stage II development of competence in *Haemophilus influenzae*, *J Bacteriol.* 101, 517-524.
- Hitchcock, P. J., Leive, L., Makela, P. H., Rietschel, E. T., Strittmatter, W. and Morrison, D. C. (1986) Lipopolysaccharide nomenclature--past, present, and future, *J Bacteriol.* 166, 699-705.
- Hoffmann, E. D., Charette, J., Stroobant, V. (1996) Mass Spectrometry: Principles and applications. West Sussex: Wiley.
- Holst, O., Brade, L., Kosma, P. and Brade, H. (1991) Structure, serological specificity, and synthesis of artificial glycoconjugates representing the genus-specific lipopolysaccharide epitope of *Chlamydia spp.*, *J Bacteriol.* 173, 1862-1866.
- Holst, O., Ulmer, A. J., Brade, H., Flad, H. D. and Rietschel, E. T. (1996) Biochemistry and cell biology of bacterial endotoxins, *FEMS Immunol Med Microbiol.* 16, 83-104.
- Hong, W., Pang, B., West-Barnette, S. and Swords, W. E. (2007) Phosphorylcholine expression by nontypeable *Haemophilus influenzae* correlates with maturation of biofilm communities in vitro and in vivo, *J Bacteriol.* 189, 8300-8307.
- Hood, D. W., Cox, A. D., Gilbert, M., Makepeace, K., Walsh, S., Deadman, M. E., Cody, A., Martin, A., Månsson, M., Schweda, E. K., Brisson, J. R., Richards, J. C., Moxon, E. R. and Wakarchuk, W. W. (2001) Identification of a lipopolysaccharide alpha-2,3-sialyltransferase from *Haemophilus influenzae*, *Mol Microbiol.* 39, 341-350.

- Hood, D. W., Cox, A. D., Wakarchuk, W. W., Schur, M., Schweda, E. K., Walsh, S. L., Deadman, M. E., Martin, A., Moxon, E. R., Richards, J. C. (2001) Genetic basis for expression of the major globotetraose-containing lipopolysaccharide from *H. influenzae* strain Rd (RM118), *Glycobiology*. 11, 957-967.
- Hood, D. W., Deadman, M. E., Allen, T., Masoud, H., Martin, A., Brisson, J. R., Fleischmann, R., Venter, J. C., Richards, J. C. and Moxon, E. R. (1996a) Use of the complete genome sequence information of *Haemophilus influenzae* strain Rd to investigate lipopolysaccharide biosynthesis, *Mol Microbiol*. 22, 951-965.
- Hood, D. W., Deadman, M. E., Jennings, M. P., Bisercic, M., Fleischmann, R. D., Venter, J. C. and Moxon, E. R. (1996b) DNA repeats identify novel virulence genes in *Haemophilus influenzae*, *Proc Natl Acad Sci U S A*. 93, 11121-11125.
- Hood, D. W., Makepeace, K., Deadman, M. E., Rest, R. F., Thibault, P., Martin, A., Richards, J. C. and Moxon, E. R. (1999) Sialic acid in the lipopolysaccharide of *Haemophilus influenzae*: strain distribution, influence on serum resistance and structural characterization, *Mol Microbiol*. 33, 679-692.
- Hood, D. W., Randle, G., Cox, A. D., Makepeace, K., Li, J., Schweda, E. K., Richards, J. C. and Moxon, E. R. (2004) Biosynthesis of cryptic lipopolysaccharide glycoforms in *Haemophilus influenzae* involves a mechanism similar to that required for O-antigen synthesis, *J Bacteriol*. 186, 7429-7439.
- Houliston, R. S., Koga, M., Li, J., Jarrell, H. C., Richards, J. C., Vitiazeva, V., Schweda, E. K., Yuki, N. and Gilbert, M. (2007) A *Haemophilus influenzae* strain associated with Fisher syndrome expresses a novel disialylated ganglioside mimic, *Biochemistry*. 46, 8164-8171.
- Janson, H., Carlsson, B., Cervin, A., Forsgren, A., Magnusdottir, A. B., Lindberg, S. and Runer, T. (1999) Effects on the ciliated epithelium of protein D-producing and -nonproducing nontypeable *Haemophilus influenzae* in nasopharyngeal tissue cultures, *J Infect Dis*. 180, 737-746.
- Jarosik, G. P. and Hansen, E. J. (1994) Identification of a new locus involved in expression of *Haemophilus influenzae* type b lipooligosaccharide, *Infect Immun*. 62, 4861-4867.
- Jones, P. A., Samuels, N. M., Phillips, N. J., Munson, R. S., Jr., Bozue, J. A., Arseneau, J. A., Nichols, W. A., Zaleski, A., Gibson, B. W. and Apicella, M. A. (2002) *Haemophilus influenzae* type b strain A2 has multiple sialyltransferases involved in lipooligosaccharide sialylation, *J Biol Chem*. 277, 14598-14611.
- Jurcisek, J., Greiner, L., Watanabe, H., Zaleski, A., Apicella, M. A. and Bakaletz, L. O. (2005) Role of sialic acid and complex carbohydrate biosynthesis in biofilm formation by nontypeable *Haemophilus influenzae* in the chinchilla middle ear, *Infect Immun*. 73, 3210-3218.
- Jurcisek, J. A. and Bakaletz, L. O. (2007) Biofilms formed by nontypeable *Haemophilus influenzae* in vivo contain both double-stranded DNA and type IV pilin protein, *J Bacteriol*. 189, 3868-3875.
- Karasic, R. B., Trumpp, C. E., Gnehm, H. E., Rice, P. A. and Pelton, S. I. (1985) Modification of otitis media in chinchillas rechallenged with nontypeable *Haemophilus influenzae* and serological response to outer membrane antigens, *J Infect Dis*. 151, 273-279.
- Keeler, J., (2005) *Understanding NMR spectroscopy*. West Sussex: Wiley.
- Kilian, M., Mestecky, J. and Schrohenloher, R. E. (1979) Pathogenic species of the genus *Haemophilus* and *Streptococcus pneumoniae* produce immunoglobulin A1 protease, *Infect Immun*. 26, 143-149.
- Landerholm, M. K., Li, J., Richards, J. C., Hood, D. W., Moxon, E. R. and Schweda, E. K. (2004) Characterization of novel structural features in the lipopolysaccharide of nondisease associated nontypeable *Haemophilus influenzae*, *Eur J Biochem*. 271, 941-953.
- Li, J., Bauer, S. H., Månsson, M., Moxon, E. R., Richards, J. C. and Schweda, E. K. (2001) Glycine is a common substituent of the inner core in *Haemophilus influenzae* lipopolysaccharide, *Glycobiology*. 11, 1009-1015.
- Li, J., Dzieciatkowska, M., Hood, D. W., Cox, A. D., Schweda, E. K., Moxon, E. R. and Richards, J. C. (2007) Structural characterization of sialylated glycoforms of *H. influenzae* by electrospray mass spectrometry: fragmentation of protonated and

- sodiated O-deacylated lipopolysaccharides, *Rapid Commun Mass Spectrom.* 21, 952-960.
- Lim, D., (1998) Microbiology 2<sup>nd</sup> edition. WCB McGraw-Hill.
- Lundström, S. L., Li, J., Deadman, M. E., Hood, D. W., Moxon, E. R. and Schweda, E. K. (2008a) Structural analysis of the lipopolysaccharide from nontypeable *Haemophilus influenzae* strain R2846, *Biochemistry.* 47, 6025-6038.
- Lundström, S. L., Li, J., Mansson, M., Figueira, M., Leroy, M., Goldstein, R., Hood, D. W., Moxon, E. R., Richards, J. C. and Schweda, E. K. (2008b) Application of capillary electrophoresis mass spectrometry and liquid chromatography multiple-step tandem electrospray mass spectrometry to profile glycoform expression during *Haemophilus influenzae* pathogenesis in the chinchilla model of experimental otitis media, *Infect Immun.* 76, 3255-3267.
- Lundström, S. L., Twelkmeyer, B., Sagemark, M. K., Li, J., Richards, J. C., Hood, D. W., Moxon, E. R. and Schweda, E. K. (2007) Novel globoside-like oligosaccharide expression patterns in nontypeable *Haemophilus influenzae* lipopolysaccharide, *Febs J.* 274, 4886-4903.
- Lysenko, E. S., Gould, J., Bals, R., Wilson, J. M. and Weiser, J. N. (2000) Bacterial phosphorylcholine decreases susceptibility to the antimicrobial peptide LL-37/hCAP18 expressed in the upper respiratory tract, *Infect Immun.* 68, 1664-1671.
- Mandrell, R. E., McLaughlin, R., Aba Kwaik, Y., Lesse, A., Yamasaki, R., Gibson, B., Spinola, S. M. and Apicella, M. A. (1992) Lipooligosaccharides (LOS) of some *Haemophilus* species mimic human glycosphingolipids, and some LOS are sialylated, *Infect Immun.* 60, 1322-1328.
- Masoud, H., Martin, A., Thibault, P., Moxon, E. R. and Richards, J. C. (2003) Structure of extended lipopolysaccharide glycoforms containing two globotriose units in *Haemophilus influenzae* serotype b strain RM7004, *Biochemistry.* 42, 4463-4475.
- Masoud, H., Moxon, E. R., Martin, A., Krajcarski, D. and Richards, J. C. (1997) Structure of the variable and conserved lipopolysaccharide oligosaccharide epitopes expressed by *Haemophilus influenzae* serotype b strain Eagan, *Biochemistry.* 36, 2091-2103.
- Mikhail, I., Yildirim, H. H., Lindahl, E. C. and Schweda, E. K. (2005) Structural characterization of lipid A from nontypeable and type f *Haemophilus influenzae*: variability of fatty acid substitution, *Anal Biochem.* 340, 303-316.
- Moriyama, S., Hotomi, M., Shimada, J., Billal, D. S., Fujihara, K. and Yamanaka, N. (2009) Formation of biofilm by *Haemophilus influenzae* isolated from pediatric intractable otitis media, *Auris Nasus Larynx.* 36, 525-531.
- Moxon, E. R., Hood, D. W., Saunders, N. J., Schweda, E. K. and Richards, J. C. (2002) Functional genomics of pathogenic bacteria, *Philos Trans R Soc Lond B Biol Sci.* 357, 109-116.
- Moxon, E. R. and Kroll, J. S. (1990a) The role of bacterial polysaccharide capsules as virulence factors, *Curr Top Microbiol Immunol.* 150, 65-85.
- Moxon, E. R. and Rappuoli, R. (1990b) *Haemophilus influenzae* infections and whooping cough, *Lancet.* 335, 1324-1329.
- Moxon, E. R., Sweetman, W. A., Deadman, M. E., Ferguson, D. J. and Hood, D. W. (2008) *Haemophilus influenzae* biofilms: hypothesis or fact?, *Trends Microbiol.* 16, 95-100.
- Murphy, T. F. (2003) Respiratory infections caused by non-typeable *Haemophilus influenzae*, *Curr Opin Infect Dis.* 16, 129-134.
- Murphy, T. F. (2009) Nontypeable *Haemophilus influenzae* as a pathogen in children, *Pediatr. Infect. Dis J.* 28, 43-48.
- Musser, J. M., Barenkamp, S. J., Granoff, D. M. and Selander, R. K. (1986) Genetic relationships of serologically nontypable and serotype b strains of *Haemophilus influenzae*, *Infect Immun.* 52, 183-191.
- Månsson, M., Bauer, S. H., Hood, D. W., Richards, J. C., Moxon, E. R. and Schweda, E. K. (2001) A new structural type for *Haemophilus influenzae* lipopolysaccharide. Structural analysis of the lipopolysaccharide from nontypeable *Haemophilus influenzae* strain 486, *Eur J Biochem.* 268, 2148-2159.

- Månsson, M., Hood, D. W., Li, J., Richards, J. C., Moxon, E. R. and Schweda, E. K. (2002) Structural analysis of the lipopolysaccharide from nontypeable *Haemophilus influenzae* strain 1003, *Eur J Biochem.* 269, 808-818.
- Månsson, M., Hood, D. W., Moxon, E. R. and Schweda, E. K. (2003a) Structural characterization of a novel branching pattern in the lipopolysaccharide from nontypeable *Haemophilus influenzae*, *Eur J Biochem.* 270, 2979-2991.
- Månsson, M., Hood, D. W., Moxon, E. R. and Schweda, E. K. (2003b) Structural diversity in lipopolysaccharide expression in nontypeable *Haemophilus influenzae*. Identification of L-glycerol-D-manno-heptose in the outer-core region in three clinical isolates, *Eur J Biochem.* 270, 610-624.
- Nizet, V., Colina, K. F., Almquist, J. R., Rubens, C. E. and Smith, A. L. (1996) A virulent nonencapsulated *Haemophilus influenzae*, *J Infect Dis.* 173, 180-186.
- O'Neill, J. M., St Geme, J. W., 3rd, Cutter, D., Adderson, E. E., Anyanwu, J., Jacobs, R. F. and Schutze, G. E. (2003) Invasive disease due to nontypeable *Haemophilus influenzae* among children in Arkansas, *J Clin Microbiol.* 41, 3064-3069.
- Phillips, N. J., Apicella, M. A., Griffiss, J. M. and Gibson, B. W. (1992) Structural characterization of the cell surface lipooligosaccharides from a nontypable strain of *Haemophilus influenzae*, *Biochemistry.* 31, 4515-4526.
- Post, D. M., Munson, R. S., Jr., Baker, B., Zhong, H., Bozue, J. A. and Gibson, B. W. (2007) Identification of genes involved in the expression of atypical lipooligosaccharide structures from a second class of *Haemophilus ducreyi*, *Infect Immun.* 75, 113-121.
- Poulsen, K., Reinholdt, J. and Kilian, M. (1992) A comparative genetic study of serologically distinct *Haemophilus influenzae* type 1 immunoglobulin A1 proteases, *J Bacteriol.* 174, 2913-2921.
- Power, P. M., Sweetman, W. A., Gallacher, N. J., Woodhall, M. R., Kumar, G. A., Moxon, E. R. and Hood, D. W. (2009) Simple sequence repeats in *Haemophilus influenzae*, *Infect Genet Evol.* 9, 216-228.
- Rahman, M. M., Gu, X. X., Tsai, C. M., Kolli, V. S. and Carlson, R. W. (1999) The structural heterogeneity of the lipooligosaccharide (LOS) expressed by pathogenic non-typeable *Haemophilus influenzae* strain NTHi 9274, *Glycobiology.* 9, 1371-1380.
- Rao, V. K., Krasan, G. P., Hendrixson, D. R., Dawid, S. and St Geme, J. W., 3rd. (1999) Molecular determinants of the pathogenesis of disease due to non-typable *Haemophilus influenzae*, *FEMS Microbiol Rev.* 23, 99-129.
- Reddy, M. S., Bernstein, J. M., Murphy, T. F. and Faden, H. S. (1996) Binding between outer membrane proteins of nontypeable *Haemophilus influenzae* and human nasopharyngeal mucin, *Infect Immun.* 64, 1477-1479.
- Rietschel, E. T., Kirikae, T., Schade, F. U., Mamat, U., Schmidt, G., Loppnow, H., Ulmer, A. J., Zahringer, U., Seydel, U., Di Padova, F. and et al. (1994) Bacterial endotoxin: molecular relationships of structure to activity and function, *Faseb J.* 8, 217-225.
- Risberg, A., Alvelius, G. and Schweda, E. K. (1999) Structural analysis of the lipopolysaccharide oligosaccharide epitopes expressed by *Haemophilus influenzae* strain RM.118-26, *Eur J Biochem.* 265, 1067-1074.
- Ruiz, N., Kahne, D. and Silhavy, T. J. (2009) Transport of lipopolysaccharide across the cell envelope: the long road of discovery, *Nat Rev Microbiol.* 7, 677-683.
- Sawardeker, J. S., Sloneker, J. H. and Jeanes, A. (1965) Quantitative Determination of Monosaccharides as Their Alditol Acetates by Gas Liquid Chromatography, *Analytical Chemistry.* 37, 1602-1604.
- Schweda, E. K., Brisson, J. R., Alvelius, G., Martin, A., Weiser, J. N., Hood, D. W., Moxon, E. R. and Richards, J. C. (2000) Characterization of the phosphocholine-substituted oligosaccharide in lipopolysaccharides of type b *Haemophilus influenzae*, *Eur J Biochem.* 267, 3902-3913.
- Schweda, E. K., Hegedus, O. E., Borrelli, S., Lindberg, A. A., Weiser, J. N., Maskell, D. J. and Moxon, E. R. (1993) Structural studies of the saccharide part of the cell envelope lipopolysaccharide from *Haemophilus influenzae* strain AH1-3 (lic3+), *Carbohydr Res.* 246, 319-330.

- Schweda, E. K., Li, J., Moxon, E. R. and Richards, J. C. (2002) Structural analysis of lipopolysaccharide oligosaccharide epitopes expressed by non-typeable *Haemophilus influenzae* strain 176, *Carbohydr Res.* 337, 409-420.
- Schweda, E. K. and Richards, J. C. (2003) Structural profiling of short-chain lipopolysaccharides from *Haemophilus influenzae*, *Methods Mol Med.* 71, 161-183.
- Schweda, E. K., Richards, J. C., Hood, D. W. and Moxon, E. R. (2007) Expression and structural diversity of the lipopolysaccharide of *Haemophilus influenzae*: Implication in virulence, *Int J Med Microbiol.* 297, 297-306.
- Severi, E., Randle, G., Kivlin, P., Whitfield, K., Young, R., Moxon, R., Kelly, D., Hood, D. and Thomas, G. H. (2005) Sialic acid transport in *Haemophilus influenzae* is essential for lipopolysaccharide sialylation and serum resistance and is dependent on a novel tripartite ATP-independent periplasmic transporter, *Mol Microbiol.* 58, 1173-1185.
- St Geme, J. W., 3rd. (2000) The pathogenesis of nontypable *Haemophilus influenzae* otitis media, *Vaccine.* 19 Suppl 1, S41-50.
- St Geme, J. W., 3rd. (2002) Molecular and cellular determinants of non-typeable *Haemophilus influenzae* adherence and invasion, *Cell Microbiol.* 4, 191-200.
- Swords, W. E., Buscher, B. A., Ver Steeg Ii, K., Preston, A., Nichols, W. A., Weiser, J. N., Gibson, B. W. and Apicella, M. A. (2000) Non-typeable *Haemophilus influenzae* adhere to and invade human bronchial epithelial cells via an interaction of lipooligosaccharide with the PAF receptor, *Mol Microbiol.* 37, 13-27.
- Swords, W. E., Moore, M. L., Godzicki, L., Bukofzer, G., Mitten, M. J. and VonCannon, J. (2004) Sialylation of lipooligosaccharides promotes biofilm formation by nontypeable *Haemophilus influenzae*, *Infect Immun.* 72, 106-113.
- Tinnert, A. S., Månsson, M., Yildirim, H. H., Hood, D. W. and Schweda, E. K. (2005) Structural investigation of lipopolysaccharides from nontypeable *Haemophilus influenzae*: investigation of inner-core phosphoethanolamine addition in NTHi strain 981, *Carbohydr Res.* 340, 1900-1907.
- Trent, M. S. (2004) Biosynthesis, transport, and modification of lipid A, *Biochem Cell Biol.* 82, 71-86.
- Tullius, M. V., Phillips, N. J., Scheffler, N. K., Samuels, N. M., Munson Jr, R. S., Jr., Hansen, E. J., Stevens-Riley, M., Campagnari, A. A. and Gibson, B. W. (2002) The lbgAB gene cluster of *Haemophilus ducreyi* encodes a beta-1,4-galactosyltransferase and an alpha-1,6-DD-heptosyltransferase involved in lipooligosaccharide biosynthesis, *Infect Immun.* 70, 2853-2861.
- Weintraub, A. (2003) Immunology of bacterial polysaccharide antigens, *Carbohydr Res.* 338, 2539-2547.
- Weiser, J. N., Maskell, D. J., Butler, P. D., Lindberg, A. A. and Moxon, E. R. (1990) Characterization of repetitive sequences controlling phase variation of *Haemophilus influenzae* lipopolysaccharide, *J Bacteriol.* 172, 3304-3309.
- Weiser, J. N. and Pan, N. (1998a) Adaptation of *Haemophilus influenzae* to acquired and innate humoral immunity based on phase variation of lipopolysaccharide, *Mol Microbiol.* 30, 767-775.
- Weiser, J. N., Pan, N., McGowan, K. L., Musher, D., Martin, A. and Richards, J. (1998b) Phosphorylcholine on the lipopolysaccharide of *Haemophilus influenzae* contributes to persistence in the respiratory tract and sensitivity to serum killing mediated by C-reactive protein, *J Exp Med.* 187, 631-640.
- Verez-Bencomo, V., Fernandez-Santana, V., Hardy, E., Toledo, M. E., Rodriguez, M. C., Heynngnezz, L., Rodriguez, A., Baly, A., Herrera, L., Izquierdo, M., Villar, A., Valdes, Y., Cosme, K., Deler, M. L., Montane, M., Garcia, E., Ramos, A., Aguilar, A., Medina, E., Torano, G., Sosa, I., Hernandez, I., Martinez, R., Muzachio, A., Carmenates, A., Costa, L., Cardoso, F., Campa, C., Diaz, M. and Roy, R. (2004) A synthetic conjugate polysaccharide vaccine against *Haemophilus influenzae* type b, *Science.* 305, 522-525.
- Vimr, E. and Lichtensteiger, C. (2002) To sialylate, or not to sialylate: that is the question, *Trends Microbiol.* 10, 254-257.
- Vimr, E., Lichtensteiger, C. and Steenbergen, S. (2000) Sialic acid metabolism's dual function in *Haemophilus influenzae*, *Mol Microbiol.* 36, 1113-1123.

- Virji, M., Kayhty, H., Ferguson, D. J., Alexandrescu, C. and Moxon, E. R. (1992) Interactions of *Haemophilus influenzae* with human endothelial cells in vitro, *J Infect Dis.* 165 Suppl 1, S115-116.
- Virji, M., Weiser, J. N., Lindberg, A. A. and Moxon, E. R. (1990) Antigenic similarities in lipopolysaccharides of *Haemophilus* and *Neisseria* and expression of a digalactoside structure also present on human cells, *Microb Pathog.* 9, 441-450.
- Viseux, N., de Hoffmann, E. and Domon, B. (1998) Structural assignment of permethylated oligosaccharide subunits using sequential tandem mass spectrometry, *Anal Chem.* 70, 4951-4959.
- Wright, J. C., Hood, D. W., Randle, G. A., Makepeace, K., Cox, A. D., Li, J., Chalmers, R., Richards, J. C. and Moxon, E. R. (2004) *lpt6*, a gene required for addition of phosphoethanolamine to inner-core lipopolysaccharide of *Neisseria meningitidis* and *Haemophilus influenzae*, *J Bacteriol.* 186, 6970-6982.
- Yildirim, H. H., Li, J., Richards, J. C., Hood, D. W., Moxon, E. R. and Schweda, E. K. (2005a) An alternate pattern for globoside oligosaccharide expression in *Haemophilus influenzae* lipopolysaccharide: structural diversity in nontypeable strain 1124, *Biochemistry.* 44, 5207-5224.
- Yildirim, H. H., Li, J., Richards, J. C., Hood, D. W., Moxon, E. R. and Schweda, E. K. (2005b) Complex O-acetylation in non-typeable *Haemophilus influenzae* lipopolysaccharide: evidence for a novel site of O-acetylation, *Carbohydr Res.* 340, 2598-2611.